Interim Report 3

COUPLED MODE THEORY FOR ADVANCED MICROWAVE DEVICES

Prepared for:

AERONAUTICAL SYSTEMS DIVISION
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

CONTRACT AF 33(657)-8343

By: *M. C. Pease*

STANFORD RESEARCH INSTITUTE

MENLO PARK, CALIFORNIA

*SRI

STANFORD RESEARCH INSTITUTE

MENLO PARK CALIFORNIA

January 1963

Interim Report 3

COUPLED MODE THEORY FOR ADVANCED MICROWAVE DEVICES

Prepared for:

AERONAUTICAL SYSTEMS DIVISION
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

CONTRACT AF 33(657)-8343

By: M. C. Pease

SRI Project No. 4052

Approved:

Philip Rice

PHILIP RICE, MANAGER PHYSICAL ELECTRONICS LABORATORY

J. Noe

JERRE D. NOE, DIRECTOR ENGINEERING SCIENCES DIVISION

Copy No. 11

ABSTRACT

This is the third interim report of a study whose purpose is to develop and apply a generalized theory of coupled modes of propagation to the study of interactions in distributed microwave devices.

In Sec. I we develop methods for the application of group theoretic techniques to conservative linear systems that may or may not be uniform. It is shown that transmission matrices that solve a particular problem do, in fact, form a group. This group is a continuous (Lie) group and can be characterized by its infinitesimal transformations. It is shown that the appropriate infinitesimal transformations can be obtained from the system operator, $\mathbf{R}(z)$. We can then investigate whether or not a given system is reducible in the group theoretic sense. If so, the given problem can be replaced by a simpler one. In any case, the properties of the system can be discussed usefully in terms of the commutation relations of the infinitesimal transformations.

In Sec. II we continue the work begun in the previous quarter on non-uniform systems such that the derivative of $\mathbf{R}$ is expressible as a commutator of $\mathbf{R}$ with another matrix $\mathbf{A}$. We showed, last quarter, that such a system is explicitly soluble if $\mathbf{A}$ can be chosen to be constant. We show, here, that this is a special case of a more general class of explicitly souble non-uniform systems.

We have been able to show that the matrix $\mathbf{A}$ describes the behavior of the eigenvectors of $\mathbf{R}$. An explicit form for $\mathbf{A}$ is obtained and some of its properties are developed.

There is reason to hope that this type of analysis will prove of great generality and power in the study of non-uniform systems. However, a considerable amount of work remains to be done to establish the method in the generality that appears to be possible.

FOREWORD

This program was established by the Aeronautical Systems Division, Wright-Patterson Air Force Base, Ohio, at Stanford Research Institute under Contract AF 33(657)-8343 for the purpose of advancing the general understanding of microwave devices through the development and application of a general theory of coupling of modes of propagation.

At Stanford Research Institute the project supervisor is Philip Rice, and the principal investigator is Marshall C. Pease.

This is ASD Project 4156; Task 41651 the project monitor is Lawrence F. Daum, (ASRNET-3) of the Electronic Technology Laboratory (ASRNE).

CONTENTS

I INTRODUCTION

The following work is based on the thought that a conservation law defines a group. That is, if we consider all possible systems that exhibit a given conservation law, we find that their transmission matrices form a continuous (Lie) group of the requisite dimensionality.

If we are given a vector differential equation that describes a system with a known conservation law, then this equation and its boundary condition determines a trajectory through the space of group elements. Alternatively, if the group elements are suitably parametrized, then the vector differential equation and its boundary conditions determine a trajectory through the space of the parameters.

The transmission matrices with different boundary conditions form a continuous set of matrices that may be embedded in a proper subgroup of the whole group generated by the conservation law. If so, then the behavior of the system can be described within this subgroup. Certain of its properties then can be obtained by consideration of the subgroup. In addition, it may happen that this subgroup can be considered as a representation of a group that is simpler than the original group. Thus, for example, we may be able to show that a given n-dimensional system can be expressed as a representation of an appropriate two-dimensional group. The n-dimensional problem, then, can be replaced by an exactly analogous two-dimensional one.

To establish this mode of analysis, we will review in some detail the formulation of the problems that concern us and the necessary parts of the theory of Lie groups.

A. THE SYSTEM AND ITS CONSERVATION LAW

We are concerned with linearized systems that are described by a vector-differential equation of the form

$$\frac{d\mathbf{x}}{dz} = -j\mathbf{R}\mathbf{x} \qquad (1)$$

The n-dimensional column vector $\mathbf{x}$ lists the significant system variables which are functions of the single independent variable, z. For example, in a TWT, $\mathbf{x}$ may list the complex amplitudes of the circuit voltage and current, and the complex amplitudes of the RF component of space charge or longitudinal beam current and velocity. If parametric interaction occurs, we can list these variables at the various frequencies that are coupled.

The matrix $\mathbf{R}$ is an $n \times n$ matrix. If $\mathbf{R}$ is independent of z, the system is said to be uniform. If $\mathbf{R}$ is a function of z, it is non-uniform. It is the latter case that is of principal interest here, since it is these problems in particular that need more subtle methods of analysis.

In what follows we shall *not* assume that $\mathbf{R}$ is constant, except as noted.

The solution to Eq. (1) can be written as

$$\mathbf{x}(z) = \mathbf{M}(z)\,\mathbf{x}(0) \tag{2}$$

where $\mathbf{M}(z)$ is an $n \times n$ matrix that is a function of z such that

$$\mathbf{M}(0) = \mathbf{I} \tag{3}$$

where $\mathbf{I}$ is the identity matrix. The constant vector $\mathbf{x}(0)$ is the boundary condition that is assumed to be applied at $z = 0$.

If we substitute Eq. (2) in Eq. (1), and require that it be valid for all $\mathbf{x}(0)$, then we find

$$\frac{d\mathbf{M}}{dz} = -j\mathbf{R}\mathbf{M} \quad . \tag{4}$$

Equation (4) with boundary condition (3), can be taken as an alternative description of the system.

We are concerned with systems that exhibit a conservation law—conservation of net real power flow, the linearized Manley-Rowe relations, Chu's kinetic power theorem, etc. Such a conservation law can be expressed as stating the invariance of a quadratic form. That is, we can define a constant matrix $\mathbf{K}$ such that the scalar, s, given by

$$s = \mathbf{x}^\dagger \mathbf{K} \mathbf{x} \tag{5}$$

is independent of z if $\mathbf{x}$ is a solution of Eq. (1). (The $\dagger$ means the hermitian conjugate, or complex conjugate transpose. Hence $\mathbf{x}^\dagger$ is a row vector and s a scalar.) We require that $\mathbf{K}$ be nonsingular, hermitian ($\mathbf{K} = \mathbf{K}^\dagger$) and independent of z.

It is easily seen, by differentiating Eq. (5) and substituting Eq. (1), that the necessary and sufficient condition on $\mathbf{R}$ is that

$$\mathbf{K}\mathbf{R} = \mathbf{R}^\dagger \mathbf{K} \quad . \tag{6}$$

We say that $\mathbf{R}$ is "K-hermitian."

If we substitute Eq. (2) in Eq. (5) and use the boundary condition, Eq. (3), we find that the necessary and sufficient condition on $\mathbf{M}$ is that

$$\mathbf{M}^\dagger \mathbf{K} \mathbf{M} = \mathbf{K} \tag{7}$$

We call such a matrix "K-unitary."

In what follows, then, we shall be concerned with the class of all K-unitary matrices. We shall show that this is a group, which we shall call the "K-group."

As mentioned in the introduction, we consider the system equation, Eq. (4) and Eq. (3) as determining a trajectory through the designated K-group.

B. THE K-GROUP

We must, first, show that Eq. (7) does, in fact, determine a group. A group is defined in terms of a set of elements and a rule of combination. In this case, the set is the set of all matrices, $\mathbf{M}$, satisfying Eq. (7), and the operation is matrix multiplication. For the set, then, to form a group, (1) it must contain the identity, $\mathbf{I}$, (2) every element must have an inverse in the set, and (3) if $\mathbf{A}$ and $\mathbf{B}$ are in the set, so is $\mathbf{AB}$.

Clearly $\mathbf{M} = \mathbf{I}$ satisfies Eq. (7). Hence the set contains the identity.

Since $\mathbf{K}$ is to be assumed nonsingular, its determinant does not vanish. Since the determinant of the product of matrices is the product of their determinants, it follows from Eq. (7) that the determinant of $\mathbf{M}$ cannot vanish. Hence $\mathbf{M}^{-1}$ exists and so does $\mathbf{M}^{\dagger -1}$. If, now, we premultiply Eq. (7) by $\mathbf{M}^{\dagger -1}$, and postmultiply by $\mathbf{M}^{-1}$, we find that

$$\mathbf{K} = (\mathbf{M}^{-1})^{\dagger}\mathbf{K}\mathbf{M}^{-1} \quad . \tag{8}$$

Hence $\mathbf{M}^{-1}$ is a member of the set. Hence, every member of the set has an inverse that is in the set.

Finally, if $\mathbf{A}$ and $\mathbf{B}$ satisfy Eq. (7), then

$$\begin{aligned}(\mathbf{AB})^{\dagger}\mathbf{K}(\mathbf{AB}) &= \mathbf{B}^{\dagger}(\mathbf{A}^{\dagger}\mathbf{KA})\mathbf{B} \\ &= \mathbf{B}^{\dagger}\mathbf{KB} = \mathbf{K}\end{aligned} \tag{9}$$

so that $(\mathbf{AB})$ is in the set.

The set of K-unitary matrices form a group, which we shall call the "K-group."

The K-group is infinite in the sense that it contains an infinity of elements.

We shall assume that the group is parametrizable at least in the neighborhood of $\mathbf{M} = \mathbf{I}$. We can define a set of matrices $\mathbf{M}_i$, that do not depend either on $\mathbf{M}$ or z, such that we can write

$$\mathbf{M} = \mathbf{I} + \Sigma\alpha_i\mathbf{M}_i \tag{10}$$

where the α_i's are functions of $\mathbf{M}$ which go to zero continuously as $\mathbf{M}$ goes to $\mathbf{I}$. We assume that this representation is single-valued for sufficiently small α_i's.

(We are skirting around some difficulties, here. To be proper, we should show that $\mathbf{M}$ is a continuous topological group, so that we have available a proper definition of neighborhood. In addition, the single-valuedness should be investigated further. However, we shall skip over these matters.)

The set $\{\mathbf{M}_i\}$ are called the "infinitesimal transformations" of the group.

Consider, now a solution of Eq. (4) which reduces to the identity at $z = z'$. This will be of the form of Eq. (10) with the variable quantities being functions of both z and z' — *i.e.*, of position and of the position of the boundary conditions.

$$\mathbf{M}(z,z') = \mathbf{I} + \Sigma\alpha_i(z,z')\mathbf{M}_i$$

$$\alpha_i(z,z') = 0 \quad \text{at} \quad z = z' \quad . \tag{11}$$

We will define the quantities

$$\gamma_i(z') = \left.\frac{\partial\alpha_i(z,z')}{\partial z}\right|_{z=z'} \quad . \tag{12}$$

The γ_i are the coefficients of the first term in the Taylor expansion of $\mathbf{M}(z,z')$ with respect to z, about the point where $\mathbf{M} = \mathbf{I}$.

If we substitute Eq. (11) in Eq. (4), we have

$$\frac{\partial\mathbf{M}(z,z')}{\partial z} = \sum_i \frac{\partial\alpha_i(z,z')}{\partial z}\mathbf{M}_i = -jR\{I + \Sigma\alpha_i(z,z')\mathbf{M}'\} \quad .$$

We evaluate this at the point $z = z'$, when the α_i become zero and find that

$$\mathbf{R}(z') = j\sum_i\gamma_i(z')\mathbf{M}_i \quad . \tag{13}$$

Hence $\mathbf{R}(z)$ can also be expressed in terms of the set $\mathbf{M}_i$. $\mathbf{R}(z)$ does not form a group, but its properties do depend on the infinitesimal transformations of the K-group under discussion.

We note that these conclusions are equally valid whether or not $\mathbf{R}$ is constant. If the system is nonuniform, the coefficients of $\mathbf{R}$, $\{j\gamma_i\}$, are functions of z, but otherwise there is no change.

It is a theorem of Lie groups that, given appropriate conditions of single-valuedness, etc., the infinitesimal transformations determine the group completely.

From knowledge of the K-group, we will be able to determine what sets of infinitesimal transformations we wish to consider. From Eq. (13)

we can determine which of these are actually involved in the **R** describing a particular system. If we can so choose the complete set that the subset involved with the given **R** are the infinitesimal transformations of a simpler group, then the system can be studied in terms of this simpler group.

C. THE ROTATION GROUP

We will find that the groups we are concerned with are related to, although not in general identical with, the three-dimensional rotation group. This is perhaps not surprising since s, defined by Eq. (5), can be regarded as a quantity analogous to the square of the length of the vector $\mathbf{x}$. That s is invariant when $\mathbf{x}$ is replaced by $\mathbf{Mx}$ then says that the length is constant under the given operation, which may be taken as the definition of a rotation.

We call this an analogy rather than a definition, because the concept of length becomes obscure. Since $\mathbf{K}$ is, usually, not positive definite, s can be zero or negative. Hence the "length" can be zero or pure imaginary, even though the vector is non-null.

Regardless of the problems of geometric visualization, it is true that these operators can be regarded as rotational.

We will, therefore, develop the properties of the rotation group from their abstract operators.

The rotation group can be described in terms of the three matrices which can be taken as its infinitesimal transformations. One possible form for these matrices is

$$\mathbf{A}_1 = \begin{pmatrix} 0 & 0 & j \\ 0 & 0 & -1 \\ -j & 1 & 0 \end{pmatrix} \tag{14}$$

$$\mathbf{A}_2 = \begin{pmatrix} 0 & 0 & -j \\ 0 & 0 & -1 \\ j & 1 & 0 \end{pmatrix} \tag{15}$$

$$\mathbf{A}_3 = \begin{pmatrix} 0 & -j & 0 \\ j & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} . \tag{16}$$

Note that we are not saying that $\mathbf{A}_1$, $\mathbf{A}_2$, and $\mathbf{A}_3$ are a group—they are not. But they are the infinitesimal transformations, *i.e.*, the $\mathbf{M}_i$, of a continuous group of $\mathbf{M}$, defined from them according to Eq. (10).

We will not here be concerned with derivation of these matrices. This can be done by proper parametrization of the rotation group. We are concerned with the properties of the abstract group, not the particular properties of rotation operators.

It can be readily verified that these three matrices have the commutation relations

$$[\mathbf{A}_1\mathbf{A}_2] = \mathbf{A}_1\mathbf{A}_2 - \mathbf{A}_2\mathbf{A}_1 = -2\mathbf{A}_3 \tag{17}$$

$$[\mathbf{A}_1\mathbf{A}_3] = -\mathbf{A}_1 \tag{18}$$

$$[\mathbf{A}_2\mathbf{A}_3] = \mathbf{A}_2 \quad . \tag{19}$$

This is not coincidence. It is a general theorem that the set of infinitesimal transformations of any group have the property that the commutation of any two of them can be expressed in terms of the set. By forming linear combinations of the infinitesimal transformations—which is equivalent to a reparametrization by linear combinations of the parameters—the commutation relations can be varied to an extent. The particular form of the $\mathbf{A}$ in Eqs. (14), (15), and (16) were carefully chosen to give Eqs. (17), (18), and (19) which, as we shall see, are particularly convenient relations.

The commutation relations are the fundamental ones. We can, for example, change the basis, *i.e.*, the coordinate axes—of the system. If we do, the infinitesimal transformations undergo a similarity transformation. They are therefore changed. But it is easy to see that the commutation relations are invariant under such a change. Hence, the commutation relations are independent of the coordinate system.

The commutators of Eqs. (17) through (19) have remarkable properties in determining the interrelations of the significant vectors of the system.

Consider, first, an eigenvector, $\mathbf{x}_i$, of $\mathbf{A}_3$:

$$\mathbf{A}_3\mathbf{x}_i = \lambda_i\mathbf{x}_i \quad . \tag{20}$$

Then A_1x_i, if it is not the null vector, is also an eigenvector of A_3 with eigenvalue $(\lambda_i + 1)$. For, from Eq. (18):

$$\begin{aligned} A_3(A_1x_i) &= (A_1A_3 + A_1)x_i \\ &= (\lambda_i + 1)A_1x_i \quad . \end{aligned} \tag{21}$$

Hence we can construct a sequence of eigenvectors of A_3 with eigenvalues that differ by unity, that are connected by A_1.

Likewise A_2x_i, if it is not the null vector, is an eigenvector of A_3, this time with eigenvalue $(\lambda_i - 1)$. For, from Eq. (19),

$$\begin{aligned} A_3(A_2x_i) &= (A_2A_3 - A_2)x_i \\ &= (\lambda_i - 1)A_2x_i \quad . \end{aligned} \tag{22}$$

Again, we can construct a sequence of eigenvectors of A_3 with eigenvalues that differ by unity, and that are connected by A_2.

We still need to show that these two sequences are the same, within scalar factors. Before doing this, however, consider the sequence formed by A_1. The number of eigenvectors of A_3 is finite. Therefore, the sequence formed by A_1, starting with any given eigenvector of A_3, has member x_M with largest eigenvalue. That is, there must exist an x_M such that

$$A_3x_M = \lambda_Mx_M \tag{23}$$

and

$$A_1x_M = 0 \quad . \tag{24}$$

Consider, now, the sequence formed from x_M by A_2. Define

$$\begin{aligned} A_2x_M &= x_{M-1} \\ A_2x_{M-k} &= x_{M-k-1} \quad . \end{aligned} \tag{25}$$

Then, using Eqs. (25), (19), and (24)

$$\begin{aligned} \mathbf{A}_1\mathbf{x}_{N-1} &= \mathbf{A}_1\mathbf{A}_2\mathbf{x}_N \\ &= -(\mathbf{A}_2\mathbf{A}_1 - 2\mathbf{A}_3)\mathbf{x}_N \\ &= -2\lambda_N\mathbf{x}_N \quad . \end{aligned} \tag{26}$$

Likewise,

$$\begin{aligned} \mathbf{A}_1\mathbf{x}_{N-2} &= \mathbf{A}_1\mathbf{A}_2\mathbf{x}_{N-1} \\ &= (\mathbf{A}_2\mathbf{A}_1 - 2\mathbf{A}_3)\mathbf{x}_{N-1} \\ &= -(2\lambda_N\mathbf{A}_2\mathbf{x}_N + 2\lambda_{N-1}\mathbf{x}_{N-1}) \\ &= -2(\lambda_N + \lambda_{N-1})\mathbf{x}_{N-1} \quad . \end{aligned} \tag{27}$$

Let us assume that

$$\mathbf{A}_1\mathbf{x}_{N-k} = -\rho_{N-k}\mathbf{x}_{N-k+1} \tag{28}$$

where

$$\rho_{N-k} = 2(\lambda_N + \lambda_{N-1} + \ldots \lambda_{N-k+1}) \tag{29}$$

which we have shown to be true for $k = 1$ and 2. Then

$$\begin{aligned} \mathbf{A}_1\mathbf{x}_{N-k-1} &= \mathbf{A}_1\mathbf{A}_2\mathbf{x}_{N-k} \\ &= (\mathbf{A}_2\mathbf{A}_1 - 2\mathbf{A}_3)\mathbf{x}_{N-k} \\ &= -(\mathbf{A}_2\rho_{N-k}\mathbf{x}_{N-k+1} + 2\lambda_{N-k}\mathbf{x}_{N-k}) \\ &= -(\rho_{N-k} + 2\lambda_{N-k})\mathbf{x}_{N-k} \quad . \end{aligned} \tag{30}$$

Hence, $\mathbf{A}_1\mathbf{x}_{N-k-1}$ is proportional to $-\mathbf{x}_{N-k}$ with scalar factor

$$\rho_{N-k-1} = \rho_{N-k} + 2\lambda_{N-k} \tag{31}$$

Hence, the sequences generated by $\mathbf{A}_1$ and $\mathbf{A}_2$ are identical within scalar factors.

We should note the possibility of more than one sequence. It is quite possible that there are more than one. If so, then the subspace spanned by each set is an invariant subspace of the whole space, and the representation is "reducible." We shall return to this later.

We have used, so far, the fact that there must be a vector in the sequence formed by $\mathbf{A}_1$ which has the largest eigenvalue. Clearly, there must also be a smallest one. That is, there must be a value m such that

$$\mathbf{A}_3\mathbf{x}_m = \lambda_m\mathbf{x}_m \tag{32}$$

$$\mathbf{A}_2\mathbf{x}_m = 0 \quad . \tag{33}$$

But, now, consider

$$\mathbf{A}_1\mathbf{A}_2\mathbf{x}_m = \mathbf{A}_1\mathbf{0} = \mathbf{0} = -\rho_{m-1}\mathbf{x}_m \quad . \tag{34}$$

Since $\mathbf{x}_m$ is not the null vector, we must have $\rho_{m-1} = 0$.

Hence, the sequence of factors ρ_{M-k} has its two end points fixed. We have, from Eq. (28) and the above, that

$$\rho_M = \rho_{m-1} = 0 \tag{35}$$

with the ρ connected by the recursion formula, Eq. (31), where the λ are connected by

$$\lambda_{M-k} = \lambda_M - k \quad . \tag{36}$$

These conditions are sufficient to determine λ_M. The solution to Eq. (31) which vanishes at $k = M$ is

$$\begin{aligned}\rho_K &= k^2 + k(2M - 2\lambda_{M-1}) \\ &\quad + M^2 - M(2M - 2\lambda_{M-1}) \quad .\end{aligned} \tag{37}$$

If we let $k = m - 1$ and require that $\rho_{m-1} = 0$, we can solve for λ_M:

$$\lambda_M = \frac{1}{2}(M - m) \quad . \tag{38}$$

It is convenient, and conventional, to let the index run over either integral or half integral values from $(-1/2)(n - 1)$ to $(+1/2)(n - 1)$, where n is the total number of vectors in the sequence. Then $M = (1/2)(n - 1)$ and $m = (-1/2)(n - 1)$ so the $\lambda_M = (1/2)(n - 1)$ and

$$\lambda_k = k \tag{39}$$

$$\rho_k = M(M + 1) - k(k + 1) \tag{40}$$

where

$$k = -\frac{1}{2}(n - 1) \ldots -1,\ 0,\ +1,\ \ldots \frac{1}{2}(n - 1) \quad \text{if } n \text{ is odd}$$

$$= -\frac{1}{2}(n - 1) \ldots -\frac{1}{2},\ +\frac{1}{2}, \ldots \frac{1}{2}(n + 1) \text{ if } n \text{ is even} \quad .$$

Thus, from the commutation laws alone, we are able to determine a great deal of the structure of the operators.

If, then, we can express $\mathbf{R}$ in terms of operators which obey these commutation laws, we will from this fact, be able to deduce many of the properties of the system. For example, suppose that we can so express $\mathbf{R}$, and suppose the boundary condition, $\mathbf{x}_0$, is an eigenvector of $\mathbf{A}_3$. Then it is evident that the state vector must, for all z, stay within the subspace spanned by the set of vectors formed by repeatedly operating on $\mathbf{x}_0$ with $\mathbf{A}_1$ and $\mathbf{A}_2$.

We must, therefore, consider the class of operators that obey the commutation relations. These are the "representations" of the rotation group.

D. ANALYSIS OF THE K-GROUP

We have discussed in some detail the three-dimensional rotation group because it is a useful example. Not only does it itself appear, on occasion, but the methods of analysis that are effective on it can be extended to more general groups.

We do wish to emphasize, however, that, in the general case, we cannot restrict ourselves to the three-dimensional rotation groups.

Consider an R that is $n \times n$. There are, then n^2 components each of which may be complex. That R is K-hermitian restricts its form, but still leaves n^2 degrees of freedom. For example, if $K = I$ so that R is hermitian, we require that the diagonal terms be real, and the off-diagonal terms be the complex conjugates of the terms in the transposed position.

Each degree of freedom may vary with z in any prescribed fashion without destroying the K-hermitian property. Hence the most general $n \times n$ K-hermitian $R(z)$ will involve n^2 arbitrary functions of z.

To express such an $\mathbf{R}$ in terms of a set of constant infinitesimal transformations $\mathbf{M}_i$, as in Eq. (14) requires n^2 elements of $\mathbf{M}_i$ so that we can match the $n^2 \gamma_i$ terms to the prescribed functions. The three-dimensional rotation group, however, has only three infinitesimal transformations. It can, at best, be used to describe an $\mathbf{R}$ that has only three prescribed functions of z.

We must, therefore, use more complicated groups to describe systems of greater complexity.

There is a useful qualification to this statement. We can measure each mode of the system against a reference phase that varies in any prescribed manner and with an amplitude scale factor that varies in any prescribed manner. The transformation to such a measurement transforms $\mathbf{R}(z)$, and can be used (at least in principle) to remove $2n$ of the degrees of freedom of $\mathbf{R}$.

For the general case, then, we can, in principle, reduce the problem to one involving $(n^2 - 2n)$ degrees of freedom--and hence, to consideration of a group that involves $(n^2 - 2n)$ infinitesimal transformations.

The difficulties of doing this completely are, very often, formidable. For example, if $n = 2$, we find we do not need any infinitesimal transformation. However, to accomplish the reduction we must exactly solve the system. If we could do this, then the group theoretic analysis would be trivial and unnecessary anyway.

It is, however, often possible to accomplish a partial reduction without excessive difficulty. For example, in the 2×2 case, we again have four degrees of freedom, initially. We can, then, represent the 2×2 case in terms of $\mathbf{I}$, $\mathbf{A}_1$, $\mathbf{A}_2$, and $\mathbf{A}_3$, where the $\mathbf{A}$'s obey the commutation laws of Eq. (17), (18), and (19). The presence of the term in $\mathbf{I}$ prevents this from being a representation of the three-dimensional rotation group. However, this term is easily removed. In fact, we can easily remove both the $\mathbf{I}$ and the $\mathbf{A}_3$ term, as we shall see later. Hence, the problem can easily be reduced to one involving the three-dimensional rotation group.

Given, then, a representation of $\mathbf{R}$ in terms of an appropriate set of infinitesimal transformations obeying known commutation relations, it is, then, possible to deduce much of the structure of the operators by methods analogous to those of the preceding section. If, for example, there exist three operators, $\mathbf{A}_1$, $\mathbf{A}_2$, and $\mathbf{A}_3$, which obey Eqs. (17), (18), (19), then all the conclusions of the preceding section apply, even though there may be other operators included in $\mathbf{R}$ which do not commute with $\mathbf{A}_1$, $\mathbf{A}_2$, and $\mathbf{A}_3$ and whose effect, therefore, is interrelated with the effect of the $\mathbf{A}_{1,2,3}$ parts of $\mathbf{R}$.

Furthermore, we can separate, to an extent, the consequences of the individual commutation laws. Equation (18), for example, is sufficient to couple, in chain fashion, the eigenvectors of $\mathbf{A}_3$. That is, if $\mathbf{x}$ is an eigenvector of $\mathbf{A}_3$, with eigenvalue of λ, then $\mathbf{A}_1\mathbf{x}$, if it is not the null vector, is also an eigenvector of $\mathbf{A}_3$ with eigenvalue $(\lambda + 1)$. $\mathbf{A}_1$ is a "raising" operator, or, in quantum mechanical terminology, is a "creation" operator.

Likewise, if Eq. (19) is true, then $\mathbf{A}_2$ is a "lowering" or "annihilation" operator.

Hence the commutation relations provide a convenient way of obtaining the structure of the system and of describing the physical consequences of the various parts of $\mathbf{R}$.

Alternatively, and perhaps more usefully, we can use the desired structure of the infinitesimal transformation to find the appropriate representation of $\mathbf{R}$. If, for example, $\mathbf{R}$ is obtained as the perturbation of a solved $\mathbf{R}_0$, then we can use the known eigenvectors of $\mathbf{R}_0$ as the modes of the system. We can, then map the couplings between these modes that are introduced by the perturbation. We can, then, take a chain of such couplings and identify it as $\mathbf{A}_1$, and the reverse chain as $\mathbf{A}_2$. There may be several such chains, each of which is identified in similar ways. Each chain, then summarizes the effect of one particular aspect of the perturbation.

We shall, shortly discuss this process in more detail.

Thus, even though the three-dimensional rotation group is inadequate for our purpose, it does serve as a useful example and prototype. We shall, therefore, in the next section, discuss in some detail its representation.

E. REPRESENTATIONS OF THE ROTATION GROUP

1. General

The irreducible and non-equivalent representations of the rotation group are known and may be found in standard textbooks. Our problem, here, however, is to establish techniques so that we may consider the converse problem. That is, we wish to write a representation of the rotation group in a form that is suitable for its identification with components of a given $\mathbf{R}$.

The most useful way of obtaining representations seems to be their development as dyad expansions on a set of K-orthogonal, maximally normalized vectors. In the situation where the given system can be considered as the result of perturbing a system whose behavior is known—usually a uniform system—the vectors can be the eigenvectors of the unperturbed system. These vectors, then, are taken as the "modes" of the system. The perturbation is considered as introducing coupling between these modes.

We have introduced the concept of K-orthogonality before, but we will review the pertinent conclusions here.

We consider a complete set of vectors $\{\mathbf{u}_i\}$. (*Complete* means that they span the whole space, so that they include a set of linearly independent vectors in terms of which any vector can be expressed. If the space is n-dimensional, then a complete set must include exactly n vectors, if they are all linearly independent.)

We introduce the symbology on the indices $(\sim i)$ which may be read as "conjugate i." It is, in other words, a relabelling of the indices. We require that this relabelling be "one-to-one," so that to each vector $\mathbf{u}_i$ there corresponds one and only one vector $\mathbf{u}_{\sim i}$, which may be the same vector, and vice versa.

The set is said to be "K-orthogonal" if and only if

$$\mathbf{u}_i^{\dagger}\mathbf{K}\mathbf{u}_{\sim j} = 0 \quad \text{if } i \neq j \quad .$$

(Note: the concept of conjugate indices is necessary only because $\mathbf{K}$ may not be positive definite. If $\mathbf{K}$ is positive definite, then we must have $\sim i$ identical with i. It is only the indefiniteness of $\mathbf{K}$ that permits the possibility of $\sim i$ being different from i.)

The set is said to be "maximally normalized" as well if

$$\mathbf{u}_i{}^\dagger \mathbf{K} \mathbf{u}_{\sim i} = \sigma_i$$

where $\sigma_i = 1$ if $i \neq \sim i$ and is ± 1 if $i = \sim i$. In the latter case, σ_i indicates the "parity" of the vector.

We have shown that, if ϕ is a coupling, or perturbation parameter, such that $R(\phi, z)$ is K-hermitian for all z, and ϕ, then the eigenvectors of the uncoupled system $R(\phi, z)$, can be so chosen as to be K-orthogonal and maximally normalized.

Given such a set of vectors, we can form the dyads $(\sigma_i \mathbf{u}_i \mathbf{u}_{\sim j}{}^\dagger \mathbf{K})$. We shall use these dyads to develop our representations.

Before proceeding to the more complex cases, let us consider, first, a comparatively simple one. Let

$$\mathbf{A}_1 = \alpha \sigma_i \mathbf{u}_i \mathbf{u}_{\sim j}{}^\dagger \mathbf{K} \tag{41}$$

$$\mathbf{A}_2 = -(1/\alpha) \sigma_j \mathbf{u}_j \mathbf{u}_{\sim i}{}^\dagger \mathbf{K} \tag{42}$$

where $i \neq j$ so that the vectors are not cross-conjugated with each other, and where α is any constant, real or complex.

Then, from Eq. (17), we find that

$$\mathbf{A}_3 = \frac{1}{2}(\sigma_i \mathbf{u}_i \mathbf{u}_{\sim i}{}^\dagger \mathbf{K} - \sigma_j \mathbf{u}_j \mathbf{u}_{\sim j}{}^\dagger \mathbf{K}) \tag{43}$$

It is easy to verify, then, that Eqs. (18) and (19) are obeyed.

It is of interest to note that the eigenvectors of $\mathbf{A}_3$ are $\mathbf{u}_i$ and $\mathbf{u}_j$ with eigenvalues $+\frac{1}{2}$ and $-\frac{1}{2}$, respectively. We also note that

$$\begin{aligned} \mathbf{A}_1 \mathbf{u}_i &= 0 & \mathbf{A}_2 \mathbf{u}_i &= -(1/\alpha)\sigma_i \sigma_j \mathbf{u}_j \\ \mathbf{A}_1 \mathbf{u}_j &= \alpha \sigma_j \sigma_i \mathbf{u}_i & \mathbf{A}_2 \mathbf{u}_j &= 0 \end{aligned} \tag{44}$$

in verification of our previous results.

This representation is then isomorphic to the representation of the rotation group as 2×2 matrices:

$$\mathbf{A}_1 = \begin{pmatrix} 0 & j \\ 0 & 0 \end{pmatrix}$$

$$\mathbf{A}_2 = \begin{pmatrix} 0 & 0 \\ j & 0 \end{pmatrix}$$

$$\mathbf{A}_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{45}$$

Therefore, if the given $\mathbf{R}$ can be expressed in terms of the operators of Eqs. (41), (42), and (43), then the problem can be replaced by an equivalent one involving the 2×2 matrices of Eq. (45).

As the next simplest situation, it may be easily verified that representations of the type of Eqs. (41) to (43) can be added if there are no cross-coupling terms. We can take, for example,

$$\mathbf{A}_1 = \alpha_i \sigma_i \mathbf{u}_i \mathbf{u}_{\sim j}{}^{\dagger}\mathbf{K} + \alpha_h \sigma_h \mathbf{u}_h \mathbf{u}_{\sim k}{}^{\dagger}\mathbf{K}$$

$$\mathbf{A}_2 = -(1/\alpha_i)\sigma_j \mathbf{u}_j \mathbf{u}_{\sim i}{}^{\dagger}\mathbf{K} - (1/\alpha_h)\sigma_k \mathbf{u}_k \mathbf{u}_{\sim h}{}^{\dagger}\mathbf{K}$$

$$\mathbf{A}_3 = \frac{1}{2}(\sigma_i \mathbf{u}_i \mathbf{u}_{\sim i}{}^{\dagger}\mathbf{K} - \sigma_j \mathbf{u}_j \mathbf{u}_{\sim j}{}^{\dagger}\mathbf{K} + \sigma_h \mathbf{u}_h \mathbf{u}_{\sim h}{}^{\dagger}\mathbf{K} - \sigma_k \mathbf{u}_k \mathbf{u}_{\sim k}{}^{\dagger}\mathbf{K}) . \tag{46}$$

We find that these matrices satisfy Eqs. (17), (18), and (19) providing i, j, k, and h are all distinct. It is interesting to note, that there is no requirement that, for example $\mathbf{u}_h$ should not be identical with $\mathbf{u}_{\sim i}$ or $\mathbf{u}_{\sim j}$. In this case, there might appear to be coupling between the two sets of vectors, but in fact, there is not.

We observe that this representation is simply the combination of two representations of the form of Eqs. (41), (42), and (43).

The eigenvectors of $\mathbf{A}_3$ are $\mathbf{u}_i$, $\mathbf{u}_j$, $\mathbf{u}_h$, and $\mathbf{u}_k$. If the $\mathbf{R}$ of a given system can be expressed in terms of $\mathbf{A}_1$, $\mathbf{A}_2$, and $\mathbf{A}_3$, then the subspace spanned by $\mathbf{u}_i$ and $\mathbf{u}_j$, and that spanned by $\mathbf{u}_h$ and $\mathbf{u}_k$, are invariant subspaces of the system. Each subspace can, then, be replaced by an analogous 2×2 problem involving the $\mathbf{A}$'s of Eq. (45).

Clearly, we can add as many representations together as we wish, providing they are wholly decoupled.

Let us consider, now a representation of higher order, We wish $\mathbf{A}_1$ to be an operator which establishes a sequence of vectors. Consider for example:

$$\mathbf{A}_1 = \alpha_1\sigma_i\mathbf{u}_i\mathbf{u}_{\sim j}{}^{\dagger}\mathbf{K} + \alpha_2\sigma_j\mathbf{u}_j\mathbf{u}_{\sim k}{}^{\dagger}\mathbf{K}$$

where $i \neq j \neq k$. Then

$$\mathbf{A}_1\mathbf{u}_i = 0$$

$$\mathbf{A}_1\mathbf{u}_j = \alpha_1\sigma_i\sigma_j\mathbf{u}_i$$

$$\mathbf{A}_1\mathbf{u}_k = \alpha_2\sigma_j\sigma_k\mathbf{u}_j$$

so that $\mathbf{A}_1$ establishes the sequence $k \to j \to i$.

We wish $\mathbf{A}_2$ to establish the same sequence in reverse order. We therefore let

$$\mathbf{A}_2 = \gamma_1\sigma_j\mathbf{u}_j\mathbf{u}_{\sim i}{}^{\dagger}\mathbf{K} + \gamma_2\sigma_k\mathbf{u}_k\mathbf{u}_{\sim j}{}^{\dagger}\mathbf{K} \tag{47}$$

so that

$$\mathbf{A}_2\mathbf{u}_i = \gamma_1\sigma_i\sigma_j\mathbf{u}_j$$

$$\mathbf{A}_2\mathbf{u}_j = \gamma_2\sigma_j\sigma_k\mathbf{u}_k$$

$$\mathbf{A}_2\mathbf{u}_k = 0 \quad . \tag{48}$$

We still need to determine the relationship between the constants. From Eq. (17) we find

$$\mathbf{A}_3 = \frac{1}{2}\{\alpha_1\gamma_1\sigma_j\mathbf{u}_j\mathbf{u}_{\sim j}{}^{\dagger}\mathbf{K} + \alpha_2\gamma_2\sigma_k\mathbf{u}_k\mathbf{u}_{\sim k}{}^{\dagger}\mathbf{K}$$

$$- \alpha_1\gamma_1\sigma_i\mathbf{u}_i\mathbf{u}_{\sim i}{}^{\dagger}\mathbf{K} - \alpha_2\gamma_2\sigma_j\mathbf{u}_j\mathbf{u}_{\sim j}{}^{\dagger}\mathbf{K}\} \quad . \tag{49}$$

We can easily find that

$$[\mathbf{A}_1\mathbf{A}_3] = \alpha_1(\alpha_1\gamma_1 - \frac{1}{2}\alpha_2\gamma_2)\sigma_i\mathbf{u}_i\mathbf{u}_{\sim j}{}^{\dagger}\mathbf{K} + \alpha_2(\alpha_2\gamma_2 - \frac{1}{2}\alpha_1\gamma_1)\sigma_j\mathbf{u}_j\mathbf{u}_{\sim k}{}^{\dagger}\mathbf{K} \quad . \tag{50}$$

For this to equal $-\mathbf{A}_1$ according to Eq. (18), we must have

$$\alpha_1\gamma_1 - \frac{1}{2}\alpha_2\gamma_2 = -1$$

$$-\frac{1}{2}\alpha_1\gamma_1 + \alpha_2\gamma_2 = -1 \quad . \tag{51}$$

This requires that

$$\alpha_1\gamma_1 = \alpha_2\gamma_2 = -2 \quad .$$

Hence we take

$$\mathbf{A}_1 = \alpha_1\sigma_i\mathbf{u}_i\mathbf{u}_{\sim j}{}^{\dagger}\mathbf{K} + \alpha_2\sigma_j\mathbf{u}_j\mathbf{u}_{\sim k}{}^{\dagger}\mathbf{K} \tag{52}$$

$$\mathbf{A}_2 = -(2/\alpha_1)\sigma_j\mathbf{u}_j\mathbf{u}_{\sim i}{}^{\dagger}\mathbf{K} - (2/\alpha_2)\sigma_k\mathbf{u}_k\mathbf{u}_{\sim j}{}^{\dagger}\mathbf{K} \tag{53}$$

$$\mathbf{A}_3 = \sigma_i\mathbf{u}_i\mathbf{u}_{\sim i}{}^{\dagger}\mathbf{K} - \sigma_k\mathbf{u}_k\mathbf{u}_{\sim k}{}^{\dagger}\mathbf{K} \quad . \tag{54}$$

It can be easily confirmed that Eq. (19) is satisfied.

The eigenvectors of $\mathbf{A}_3$ are $\mathbf{u}_i$, $\mathbf{u}_j$, and $\mathbf{u}_k$ with eigenvalues +1, 0, −1, as expected.

This representation is isomorphic to the 3 × 3 representation of the rotation group, for which we may use the matrices given in Eqs. (14), (15), and (16).

Clearly, again, we can add representations together providing they do not have vectors in common.

The same technique can be used to develop representation coupling any number of vectors together. We take $\mathbf{A}_1$ as the sum of scalars times the dyads that move us along the chosen sequence. We take $\mathbf{A}_2$ as the sum of different scalars times the dyads that move us along the same sequence in the reverse order. We can, then, compute $\mathbf{A}_3$ in general terms from Eq. (17). From either Eq. (17), or (18) we obtain the necessary relations between the scalars of $\mathbf{A}_1$ and those of $\mathbf{A}_2$. This, then, fixes $\mathbf{A}_1$, $\mathbf{A}_2$, and $\mathbf{A}_3$, within the allowable variations, so that Eqs. (17), (18), and (19) are satisfied. We have, then, a suitable representation of the rotation group.

In this way, we are able to obtain a variety of representations of the rotation group, and to discuss them in terms of the coupling of vectors.

2. Example

To illustrate our discussion of the meaning of the different representations, we shall consider a simple case.

Let us consider the system matrix of a transmission line:

$$\mathbf{R} = \begin{pmatrix} 0 & \beta/Z \\ \beta/Z & 0 \end{pmatrix} \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{55}$$

for which a suitable K is

$$\mathbf{K} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} . \tag{56}$$

We are not here assuming that Z is independent of z so that this may, itself, be a non-trivial problem. We shall, however, show the relation between this problem and one of higher dimensionality, and therefore greater apparent complexity.

If we take

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \mathbf{u}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{57}$$

then $1 \sim 2$ and $\sigma = 1$. Hence

$$\begin{aligned} \mathbf{A}_1 &= \alpha \mathbf{u}_1 \mathbf{u}_1^{\dagger} \mathbf{K} = \alpha \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \\ \mathbf{A}_2 &= -\frac{1}{\alpha} \mathbf{u}_2 \mathbf{u}_2^{\dagger} \mathbf{K} = \frac{1}{\alpha} \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} \\ \mathbf{A}_3 &= \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \end{aligned} \tag{58}$$

Hence

$$\mathbf{R} = \frac{\beta Z}{\alpha}\mathbf{A}_1 - \frac{\beta\alpha}{Z}\mathbf{A}_2 \quad . \tag{59}$$

What we want to do now is to consider whether or not there are situations of greater apparent complexity which are, nevertheless, essentially the same as Eq. (55), and for which a solution to Eq. (55) can be used. This will be true if Eq. (59) is true, but with an $\mathbf{A}_1$ and $\mathbf{A}_2$ that is more complex than Eq. (58).

To consider this problem, let us proceed in the opposite direction by considering a more complex vector constructed from the vector of Eq. (55). Let us consider:

$$\mathbf{y} = \begin{pmatrix} ax_1^2 \\ bx_1x_2 \\ cx_2^2 \end{pmatrix} \tag{60}$$

where a, b, and c are constants to be determined. This form is a generalization of the "direct product" or Kronecker product of $\mathbf{y}$ with itself. We see, then, that

$$\frac{d\mathbf{y}}{dz} = \begin{pmatrix} 2ax_1(dx_1/dz) \\ bx_1(dx_2/dz) + bx_2(dx_1/dz) \\ 2cx_2(dx_2/dz) \end{pmatrix}$$

$$= \begin{pmatrix} 2ax_1(-j\beta Zx_2) \\ bx_1(-j\beta x_1/Z) + bx_2(-j\beta Zx_2) \\ + 2cx_2(-j\beta x_1/Z) \end{pmatrix}$$

$$= -j\begin{pmatrix} 0 & 2a\beta Z/b & 0 \\ b\beta/aZ & 0 & b\beta Z/c \\ 0 & 2c\beta/cZ & 0 \end{pmatrix}\begin{pmatrix} ax_1^2 \\ bx_1x_2 \\ cx_2^2 \end{pmatrix} \quad . \tag{61}$$

The conservation law states that

$$s = \mathbf{x}^\dagger\mathbf{K}\mathbf{x} = x_1^*x_2 + x_1x_2^* \tag{62}$$

is invariant. We require the existence of a conservation law for the **y** vector. We can obtain such a law by observing that

$$s^2 = x_1{}^{*2}x_2{}^2 + 2x_1{}^*x_2{}^*x_1 + x_2{}^{*2}x_1{}^2 \tag{63}$$

must also be invariant.

For this to be a conservation law for the **y** vector we require, within a constant factor, that

$$\begin{aligned} a &= e^{j\psi}/\sqrt{A}^* \\ b &= \sqrt{2} \\ c &= e^{j\psi}/\sqrt{A} \quad . \end{aligned} \tag{64}$$

Then, if we set

$$\mathbf{K}' = \begin{pmatrix} 0 & 0 & A \\ 0 & 1 & 0 \\ A^* & 0 & 0 \end{pmatrix} \tag{65}$$

and

$$\mathbf{R}' = \begin{pmatrix} 0 & \sqrt{2}e^{j\psi}\beta Z/\sqrt{A^*} & \\ \sqrt{2}A^*e^{-j\psi}\beta/Z & 0 & \sqrt{2}Ae^{-j\psi}\beta Z \\ 0 & \sqrt{2}e^{j\psi}\beta/Z\sqrt{A} & 0 \end{pmatrix} \tag{66}$$

we can easily see that $\mathbf{R}'$ is K'-hermitian.

We can, now, set

$$\mathbf{A}_1' = \sqrt{2}\begin{pmatrix} 0 & \alpha_1 & 0 \\ 0 & 0 & \alpha_2 \\ 0 & 0 & 0 \end{pmatrix} .$$

$$\mathbf{A}' = \sqrt{2}\begin{pmatrix} 0 & 0 & 0 \\ -1/\alpha_1 & 0 & 0 \\ 0 & -1/\alpha_2 & 0 \end{pmatrix}$$

$$\mathbf{A}_3' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} . \tag{67}$$

and find that these matrices obey the conservation rules.

If we set

$$\alpha_1 = \alpha e^{j\psi}/\sqrt{A}^*$$

$$\alpha_2 = \alpha e^{-j\psi}\sqrt{A} \tag{68}$$

then we find that

$$\mathbf{R}' = \frac{\beta Z}{\alpha}\mathbf{A}_1' - \frac{\beta\alpha}{Z}\mathbf{A}_2' \tag{69}$$

which is the same form as Eq. (59).

The point of this development is not the determination of $\mathbf{R}'$ from $\mathbf{R}$, as we have done, but rather the converse.

Given a three-dimensional system with the matrix $\mathbf{R}'$ of the form of Eq. (66), where Z and β may be functions of z. We do not need to solve it directly. Instead, it is sufficient to solve the equivalent two-dimensional system with the $\mathbf{R}$ of Eq. (55). The solution of the three-dimensional system is, then, the vector $\mathbf{y}$ which is determined through Eq. (60), from $\mathbf{x}$, the solution of the two-dimensional problem.

This technique, then offers hope that a multidimensional problem can be reduced to one or a sequence of problems of lower dimensionality and the answer obtained by the appropriate direct products of the solutions of the simpler problems.

F. CONSTANT $\mathbf{R}$

Since a uniform system (*i.e.*, one such that $\mathbf{R}$ is of simple structure and constant) is, in principle, always soluble by the eigenvectors (and, if necessary, the generalized eigenvectors) of $\mathbf{R}$, the group theoretic properties of such systems are principally of academic interest. Consideration of such systems, however, may clarify the significance of some of the statements that we have made.

There are several ways of stating the property that makes uniform systems directly soluble. We may observe, first, that the eigenvectors themselves each form a one-dimensional representation of the rotation group. This is the trivial representation of the rotation group that assigns to each operator the scalar 1.

In more conventional terms, we know that we can use the eigenvectors as a basis so that $\mathbf{R}$ becomes diagonal. The appropriate similarity transformation is a constant one, and the system has, then, been resolved into K-orthogonal and maximally normalized modes. It is these modes that permit direct solution.

A third way of putting it is to say that the eigenvectors each generates a one-dimensional subspace that is invariant under $\mathbf{R}$ and therefore under any $\mathbf{M}$ that is a solution of the system.

When we look for two-dimensional representations, we are in fact looking for two-dimensional subspaces that are invariant under $\mathbf{R}$. Then, if $\mathbf{x}$ initially lies in such a subspace, it continuously transforms into a vector in the subspace, and hence is retained permanently within it. The subspace, again, is invariant under the system matrix, $\mathbf{M}$.

A change to such a basis transforms $\mathbf{R}$ into a quasi-diagonal form with a 2×2 matrix on the diagonal in the position corresponding to the given subspace. If the basis is such that the entire space is divided into disjoint invariant subspaces, the transformed $\mathbf{R}$ is quasi-diagonal with 2×2 matrices along the diagonal.

Any subspace that is spanned by any two eigenvectors of $\mathbf{R}$ is evidently, an invariant subspace. On the other hand, if the eigenvectors are not degenerate, such a subspace can be identified by two vectors that are not eigenvectors of $\mathbf{R}$.

It is not evident that any two-dimensional invariant subspace is necessarily a representation of the rotation group. One finds, generally, an appropriate determination of $\mathbf{A}_1$ and $\mathbf{A}_2$. However, there is no guarantee that the resultant $\mathbf{A}_3$, determined from Eq. (17), will "fit" the remaining terms. If not, however, we can correct the discrepancy.

It is shown in the appendix that we can make a transformation of $\mathbf{x}$, and hence of $\mathbf{R}$, that eliminates all "self-coupling" terms—*i.e.,* all terms involving $(\sigma_i \mathbf{u}_i \mathbf{u}_{\sim i}{}^{\dagger}\mathbf{K})$. Hence we can, after such a transformation, always write any two-dimensional subspace in terms of $\mathbf{A}_1$ and $\mathbf{A}_2$ alone.

The transformations that are involved amount to a z-dependent renormalization of the modes of the system. Specifically, we consider the system that would occur if there were no cross-coupling terms—*i.e.,* if all terms involved $\sigma_i \mathbf{u}_i \mathbf{u}_{\sim j}{}^{\dagger}\mathbf{K}(i \neq j)$ were eliminated. The solution to this system is used as the reference for the full system, the phase angles and amplitudes of the modes being measured by reference to the solution of the non-cross-coupled system.

This transformation is valid not only for the uniform systems considered in this section, but also for non-uniform systems. It is always possible to eliminate completely all self-coupling terms, regardless of their z-dependence.

G. NON-UNIFORM SYSTEMS

If $\mathbf{R}$ is a function of z, then we have a different situation.

If the z-dependence is only in the eigenvalues of $\mathbf{R}$, then the eigenvectors can be chosen to be constant, and the previous considerations apply without exception.

If, however, it is not possible to choose eigenvectors that are constant, then we can no longer use the one-dimensional representations.

In general, in this case, we must operate with invariant subspaces of $\mathbf{R}$ that are independent of z. If we attempt to use subspaces that are z-dependent, we can still find operators with the proper commutation rules, in terms of which $\mathbf{R}$ can be decomposed. However, these operators will not be independent of z, and it does not generally follow, then, that $\mathbf{M}$ can also be expressed in these terms. It will follow only if the derivatives of $\mathbf{A}_1$, $\mathbf{A}_2$, and $\mathbf{A}_3$ are also decomposable in terms of $\mathbf{A}_1$, $\mathbf{A}_2$, and $\mathbf{A}_3$. This appears to be a rather stringent condition, although it may lead to some special situations of interest.

That constant invariant subspaces exist is evident. At worst we can take the whole space, which is constant.

Of probably greater interest are those cases in which there exist proper invariant subspaces. Assuming it is possible to express $\mathbf{R}$, possibly after adjustment of the reference phases, in terms of constant $\mathbf{A}$'s within the subspace, with the z-dependence contained wholly within the γ factors, then $\mathbf{R}$ will be isomorphic two-dimensional representations of the rotation group. The problem, then, can be replaced by a simpler one.

H. DECOMPOSITION OF $d\mathbf{R}/dz$ INTO A COMMUTATOR

As an immediate application of some of these concepts, we will consider the case when $\mathbf{R}$, which is not necessarily constant but has the property that there exists a constant matrix $\mathbf{C}$ such that

$$\frac{d\mathbf{R}}{dz} = \mathbf{CR} - \mathbf{RC} = [\mathbf{C}, \mathbf{R}] \quad . \tag{70}$$

It was shown in the Second Interim Report that this class of non-uniform systems has the explicit solution

$$\mathbf{M}(z) = \exp(-j\mathbf{R} - \mathbf{C})z \exp \mathbf{C}z \quad . \tag{71}$$

A problem that was unanswered at that time was the determination of the systems that belong to this class, what their general properties are, and what tests can be made to determine if a given system belongs to the class. Also, no systematic procedure was available for the determination of the constant $\mathbf{C}$ matrix.

We shall find that the representation of $\mathbf{R}$ in terms of the infinitesimal transformations $\mathbf{A}_1$, $\mathbf{A}_2$, and $\mathbf{A}_3$ of the rotation group, when this is possible, allows us to find an answer to these questions in certain circumstances.

We shall here confine our attention to the case when the system can be expressed as a trajectory in a representation of the three-dimensional rotation group. That is, we will assume that there exists constant operators $\mathbf{A}_1$, $\mathbf{A}_2$, and $\mathbf{A}_3$ such that Eqs. (17) through (19) are obeyed and such that $\mathbf{R}$ can be written as

$$\mathbf{R} = \gamma_1\mathbf{A}_1 + \gamma_2\mathbf{A}_2 + \gamma_3\mathbf{A}_3$$

so that

$$\frac{d\mathbf{R}}{dz} = \frac{d\lambda_1}{dz}\mathbf{A}_1 + \frac{d\lambda_2}{dz}\mathbf{A}_2 + \frac{d\lambda_3}{dz}\mathbf{A}_3 \quad . \tag{72}$$

It appears likely that other, more complicated situations could be handled in similar fashion.

It is reasonable, then, to look for a constant matrix $\mathbf{C}$ which can be represented in the same way:

$$\mathbf{C} = \alpha_1 \mathbf{A}_1 + \alpha_2 \mathbf{A}_2 + \alpha_3 \mathbf{A}_3 \quad . \tag{73}$$

To $\mathbf{C}$ we can add any other matrix that commutes with $\mathbf{R}$, or that commute separately with $\mathbf{A}_1$, $\mathbf{A}_2$, and $\mathbf{A}_3$.

The commutator then is

$$\begin{aligned}
[\mathbf{C}, \mathbf{R}] &= j(\alpha_1\gamma_2 - \alpha_2\gamma_1)[\mathbf{A}_1, \mathbf{A}_2] \\
&\quad + j(\alpha_1\gamma_3 - \alpha_3\gamma_1)[\mathbf{A}_1, \mathbf{A}_3] \\
&\quad + j(\alpha_2\gamma_3 - \alpha_3\gamma_2)[\mathbf{A}_2, \mathbf{A}_3] \\
&= -2j(\alpha_1\gamma_2 - \alpha_2\gamma_3)\mathbf{A}_3 \\
&\quad -j(\alpha_1\gamma_3 - \alpha_3\gamma_1)\mathbf{A}_1 \\
&\quad +j(\alpha_2\gamma_3 - \alpha_3\gamma_2)\mathbf{A}_2 \quad .
\end{aligned} \tag{74}$$

Comparing Eqs. (72) and (74), we see that the necessary and sufficient conditions, under the given assumptions, for the existence of a decomposition into a commutator are that

$$\begin{aligned}
\frac{d\gamma_1}{dz} &= -\alpha_1\gamma_3 + \alpha_3\gamma_1 \\
\frac{d\gamma_2}{dz} &= \alpha_2\gamma_3 - \alpha_3\gamma_2 \\
\frac{d\gamma_3}{dz} &= -2\alpha_1\gamma_2 + 2\alpha_2\gamma_1
\end{aligned} \tag{75}$$

or that

$$\frac{d}{dz}\begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix} = \begin{pmatrix} \alpha_3 & 0 & -\alpha_1 \\ 0 & -\alpha_3 & \alpha_2 \\ 2\alpha_2 & -2\alpha_1 & 0 \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix} \quad . \tag{76}$$

If the γ are solutions of Eq. (76), and then Eq. (70) holds. This does not depend on the constancy of the α, and therefore is the general condition for Eq. (70).

If the α are required to be constant, then Eq. (76) has the solution

$$\boldsymbol{\Gamma}(z) = \exp(\mathbf{P}z)\boldsymbol{\Gamma}(0) \tag{77}$$

where

$$\boldsymbol{\Gamma} = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix} \tag{78}$$

and $\mathbf{P}$ is the matrix of Eq. (76).

We can find, then, by the usual methods that

$$e^{\mathbf{P}z} = \kappa^2 \begin{pmatrix} \begin{pmatrix} 2\alpha_1\alpha_2 + \alpha_3\kappa \sin \kappa z \\ +(2\alpha_1\alpha_2 - \kappa^2) \cos \kappa z \end{pmatrix} & 2\alpha_1^2(1 - \cos \kappa z) & \begin{pmatrix} -\alpha_1\kappa \sin \kappa z \\ -\alpha_1\alpha_3(1 - \cos \kappa z) \end{pmatrix} \\ 2\alpha_2^2(1 - \cos \kappa z) & \begin{pmatrix} 2\alpha_1\alpha_2 - \alpha_3\kappa \sin \kappa z \\ +(2\alpha_1\alpha_2 - \kappa^2) \cos \kappa z \end{pmatrix} & \begin{pmatrix} \alpha_2\kappa \sin \kappa z \\ -\alpha_2\alpha_3(1 - \cos \kappa z \end{pmatrix} \\ \begin{pmatrix} 2\alpha_2\kappa \sin \kappa z \\ +2\alpha_2\alpha_3(1 - \cos \kappa z) \end{pmatrix} & \begin{pmatrix} -2\alpha_1\kappa \sin \kappa z \\ +2\alpha_1\alpha_3(1 - \cos \kappa z) \end{pmatrix} & \begin{pmatrix} \kappa^2 - 4\alpha_1\alpha_2 \\ +4\alpha_1\alpha_2 \cos \kappa z \end{pmatrix} \end{pmatrix} \tag{79}$$

where

$$\kappa^2 = 4\alpha_1\alpha_2 - \alpha_3^2 \tag{80}$$

and may be real, imaginary, or complex.

This is valid providing $\kappa \neq 0$. If $\kappa = 0$, the solution is

$$e^{Pz} = \begin{pmatrix} (1 + \alpha_3 z/2)^2 & \alpha_1^2 z^2 & -\alpha_1 z(1 + \alpha_3 z/2) \\ \alpha_2^2 z^2 & (1 - \alpha_3 z/2)^2 & \alpha_2 z(1 - \alpha_3 z/2) \\ 2\alpha_2 z(1 + \alpha_3 z/2) & -2\alpha_1 z(1 - \alpha_3 z/2) & 1 - \alpha_3^2 z^2/2 \end{pmatrix} . \tag{81}$$

These solutions are easily obtained by considering the power series expansion of the exponential and using the Cayley-Hamilton Theorem which, in this case, tells us that

$$\mathbf{P}^3 = -\kappa^2\mathbf{P} \quad .$$

Therefore, a necessary and sufficient condition under the given assumption is that the γ vary according to Eq. (77) where $\exp(\mathbf{P}z)$ has the form either of Eq. (79) or (81). Such a set of γ generates a class of systems that can be solved by the method used here. Furthermore, the matching to Eq. (79) or (81) will determine the α, and therefore the $\mathbf{A}$.

As a test for whether a given system belongs to this class, the formulas given leave something to be desired. We can develop some tests which are necessary, although not sufficient, and which are far easier to apply.

We can consider Eq. (75) as determining the α in terms of the γ. We observe that the determinant of the right-hand side vanishes. Therefore these equations must be redundant. We can find that the condition for this is that

$$2\gamma_1 \frac{d\gamma_2}{dz} - \gamma_3 \frac{d\gamma_3}{dz} + 2\gamma_2 \frac{d\gamma_1}{dz} = 0 \tag{82}$$

or that

$$4\gamma_1\gamma_2 = \gamma_3^2 + c \tag{83}$$

where c is a constant.

Equation (83) is a necessary and sufficient condition, if $\mathbf{R}$ and $\mathbf{C}$ can be represented in terms of the rotation group, for $d\mathbf{R}/dz$ to be expressible as a commutator at all, although it is not sufficient to assure the constancy of $\mathbf{C}$. We presume that Eq. (83) is, in some way, equivalent to the condition that the eigenvalues of $\mathbf{R}$ be constant.

We can also eliminate from Eqs. (75) any two of the γ. We find then, assuming that the α are constant,

$$(d^3/dz^3 - \alpha_3^2 d/dz - 4\alpha_1\alpha_2\alpha_3)\gamma_i = 0 \quad . \tag{84}$$

It is necessary, although not sufficient, that each of the γ shall obey an equation of the form of Eq. (84) with the same coefficients for a constant **C** to exist.

II MATHEMATICAL TECHNIQUES FOR NON-UNIFORM COUPLED-MODE SYSTEMS

A. INTRODUCTION

In Part II-A of Interim Report 2, we considered systems described by the vector differential equation

$$\frac{d\mathbf{x}(z)}{dz} = -j\mathbf{R}(z)\mathbf{x}(z) \tag{85}$$

where **R** may be a function of z. We considered, in particular, those systems in which **R** is related to a matrix **A** by the equation

$$\frac{d\mathbf{R}}{dz} = -j[\mathbf{A}, \mathbf{R}] = -j(\mathbf{AR} - \mathbf{RA}) \tag{86}$$

square brackets being used for the commutator.

Interest in such systems was stimulated by the fact that, if **A** is constant--*i.e.*, independent of z--then the system is soluble. Thus, this is a class of non-uniform systems for which it is possible to obtain an exact and explicit solution.

Our purpose here is to study this situation in greater depth so as to obtain better understanding of what is going on, and to determine, if possible, how to generalize the class of soluble systems in some useful manner.

We may note, at the start, an interesting and suggestive symmetry. Suppose there exists a matrix, **K**, which we call a "metric" and which is constant, non-singular, and hermitian ($\mathbf{K} = \mathbf{K}^{\dagger}$, the dagger meaning the complex conjugate transpose). And suppose that **R** is, at all values of z, K-hermitian so that

$$\mathbf{KR} = \mathbf{R}^{\dagger}\mathbf{K} \cdot \tag{87}$$

[The general significance of this relation we have discussed elsewhere. The pertinent aspects will be reviewed later. For the moment, let us simply take Eq. (87) as given.]

Let us define the matrix $\mathbf{X}$ as

$$\mathbf{X} = \mathbf{x}\mathbf{x}^{\dagger}\mathbf{K} \quad . \tag{88}$$

Then, since $\mathbf{K}$ is constant

$$\begin{aligned}
\frac{d\mathbf{X}}{dz} &= \frac{d\mathbf{x}}{dz}\mathbf{x}^{\dagger}\mathbf{K} + \mathbf{x}\frac{d\mathbf{x}^{\dagger}}{dz}\mathbf{K} \\
&= -j\mathbf{R}\mathbf{x}\mathbf{x}^{\dagger}\mathbf{K} + j\mathbf{x}\mathbf{x}^{\dagger}\mathbf{R}^{\dagger}\mathbf{K} \\
&= -j\mathbf{R}\mathbf{x}\mathbf{x}^{\dagger}\mathbf{K} + j\mathbf{x}\mathbf{x}^{\dagger}\mathbf{K}\mathbf{R} \\
&= -j\mathbf{R}\mathbf{X} + j\mathbf{X}\mathbf{R} \\
&= -j\,[\mathbf{R}, \mathbf{X}] \quad .
\end{aligned} \tag{89}$$

Equation (89) is of the same form as Eq. (86), but with $\mathbf{X}$ in place of $\mathbf{R}$ and $\mathbf{R}$ in place of $\mathbf{A}$. We may suspect then, that Eq. (86) is a useful alternative representation of the system represented by Eq. (89). If $\mathbf{A}$ is constant, it is certainly simpler.

Even if $\mathbf{A}$ cannot be chosen as constant, it may be possible to so choose $\mathbf{A}$ that Eq. (86) is still a considerable simplification.

This line of thought suggests the possibility of continuing the process. If $\mathbf{A}$ cannot be chosen as constant, is it possible to choose it so that

$$\frac{d\mathbf{A}}{dz} = -j\,[\mathbf{B}, \mathbf{A}] \tag{90}$$

where $\mathbf{B}$ is constant? Or if not, can we choose $\mathbf{B}$ so that

$$\frac{d\mathbf{B}}{dz} = -j\,[\mathbf{C}, \mathbf{B}] \tag{91}$$

where $\mathbf{C}$ is constant? And so on.

If, at any point in this process, we do obtain a constant matrix, then we can start from that point and try to work our way back up the line.

If, for example, $\mathbf{C}$ is constant in Eq. (91), then the solution to Eq. (91) is

$$\mathbf{B} = e^{-j\mathbf{C}z}\mathbf{B}_0 e^{j\mathbf{C}z} \quad . \tag{92}$$

This expression for $\mathbf{B}$ can then be substituted in Eq. (90) and a solution for $\mathbf{A}$ sought. And so on.

This suggests that we have, here, a powerful tool for the analysis of non-uniform systems. Furthermore, it provides a classification of such systems according to the number of steps involved before a constant matrix is reached.

The results that will be reported here do not completely establish this method of analysis. There remain several important questions for which we have been unable, so far, to find rigorous answers. In particular, we may cite the following:

(1) Given a matrix $\mathbf{R}$ which can be expressed as Eq. (86). Is it always possible to choose $\mathbf{A}$ so as to obtain Eq. (90)--*i.e.*, which has constant eigenvalues? It appears that we always can, at least for $\mathbf{R}$ K-hermitian. But we have not proven it.

(2) Assuming that the answer to the first question is yes, does the sequence of Eqs. (86), (90), (91) always lead eventually to a constant matrix, or can it be made to? We suspect that it does not--that it is possible to have an $\mathbf{R}$ such that the sequence of equations cannot be made to terminate. However, this is pure speculation.

(3) Assuming that the answer to the second question is no, then in what sense is Eq. (90) simpler than Eq. (86), Eq. (91) simpler than Eq. (90), etc. Is it a valid approximation scheme, for example, to take $\mathbf{R} = \mathbf{R}_0$ as the first approximation, $\mathbf{A} = \mathbf{A}_0$ as the second, etc? We suspect that it is, but we have not proven convergence, or studied the rate of convergence to determine its usefulness.

There are, then, serious questions remaining. Nevertheless, the results we have obtained are sufficient to indicate the value of the approach.

B. K-CONSERVATIVE SYSTEMS

Before considering the problem itself, we shall review some of the pertinent implications of the existence of a $\mathbf{K}$ metric.

We restrict our attention to systems that exhibit a quadratic conservation law. That is, we assume that there exists a constant nonsingular matrix $\mathbf{K}$, which can be taken as hermitian, such that the scalar quantity, s,

$$s = \mathbf{x}^{\dagger}\mathbf{K}\mathbf{x} \tag{93}$$

is preserved (*i.e.*, is independent of z) if the vector is any solution of the system.

The conservation of energy, the linearized Manley-Rowe relations, and Chu's kinetic power theorem are examples of such a conservation law and lead to such a $\mathbf{K}$. It appears that any system that does not contain loss mechanisms has such a law, although this has not been generally proven. At least it can be said that the class of systems with such a law does include a great many systems of considerable interest.

If we differentiate Eq. (93) and substitute Eq. (85), we see that a sufficient condition is that $\mathbf{R}$ be K-hermitian, or that Eq. (87) hold. The necessity of this requires a little more subtle analysis, but can be shown, also.

We have also shown that Eq. (87) implies that the eigenvectors and, if necessary, the generalized eigenvectors of $\mathbf{R}$ can be chosen so that they are "K-orthogonal and maximally normalized." That is, if the set of vectors $\{\mathbf{u}_i\}$ are the eigenvectors and generalized eigenvectors and, if any degeneracies among them are suitably resolved, then they can be so normalized that

$$\mathbf{u}_i^{\dagger}\mathbf{K}\mathbf{u}_{\sim j} = \sigma_i \delta_{ij} \tag{94}$$

where $\sigma_i = \pm 1$ and δ_{ij} is the Kronecker delta and is 1 if $i = j$, or 0 otherwise. The symbol $(\sim j)$ which can be read "conjugate j" indicates a possible relabelling of the indices. For a given $\mathbf{u}_i$, there is only a single vector of the set such that the scalar, Eq. (94), does not vanish. We call this vector $\mathbf{u}_{\sim i}$. It may be the same as $\mathbf{u}_i$, or it may be different.

The complication of the conjugate indices, and the parity terms, σ_i, arise because $\mathbf{K}$ may not be, and in general is not, positive definite.

In the analysis that follows, we will assume that $\mathbf{R}$ is K-hermitian and that we know, or can find, a complete set of vectors $\{\mathbf{u}_i(z)\}$ that are, at all z, eigenvectors and generalized eigenvectors of $\mathbf{R}$, and that are everywhere K-orthogonal and maximally normalized, and such that the σ_i are everywhere constant.

We shall also assume that these vectors can be differentiated as many times as necessary. This, in effect, assumes not only that $\mathbf{R}$ can be differentiated but that its "structure" remains constant. If, for example, $\mathbf{R}$ includes a description of the transverse modes of an electron beam, we must require that the longitudinal field shall not reverse sign, since this would reverse the parities of the transverse modes.

C. GENERAL THEOREMS

We shall, now, prove some theorems that are of value in establishing, partially, this method of analysis, and that help to give insight into its significance.

Theorem 1—A necessary and sufficient condition for the existence of a matrix $\mathbf{V}$, constant or not, that satisfies

$$\frac{d\mathbf{U}}{dz} = -j[\mathbf{V}, \mathbf{U}] \tag{95}$$

is that the eigenvalues of $\mathbf{U}$ be constant.

This theorem was proven in Part II-A of Interim Report 2, and we shall not repeat its proof here.

If we consider $\mathbf{U} = \mathbf{R}$, this appears to be a restriction on the system being considered. It may not be, however. If, in Eq. (85), we let

$$\mathbf{x} = \mathbf{S}\mathbf{y}$$

where $\mathbf{S}$ is a function of z, then

$$\frac{d\mathbf{y}}{dz} = -j\,\mathbf{S}^{-1}\mathbf{R}\mathbf{S} - j\mathbf{S}^{-1}\frac{d\mathbf{S}}{dz}\mathbf{y} \tag{96}$$

and the question now is whether we can so choose $\mathbf{S}$ that

$$\mathbf{R}' = \mathbf{S}^{-1}\mathbf{R}\mathbf{S} - j\mathbf{S}^{-1}\frac{d\mathbf{S}}{dz} \tag{97}$$

has constant eigenvalues. We suspect that this is always possible, at least in principle. If so, then the entire class of K-conservative non-uniform systems can be studied. This is, however, only a conjecture at the present time.

We now seek a relation between the $\mathbf{V}$ of Eq. (95) and the detailed structure of $\mathbf{U}$. It is provided by the following theorem:

Theorem 2—If we have a complete set of $\{\mathbf{u}_i\}$ of eigenvectors and generalized eigenvectors of $\mathbf{U}$ such that

$$\frac{d\mathbf{u}_i}{dz} = -j\mathbf{V}\mathbf{u}_i \tag{98}$$

then Eq. (95) is satisfied with this matrix $\mathbf{V}$. Conversely, given Eq. (95), we can choose a complete set of eigenvectors and generalized eigenvectors of $\mathbf{U}$ that satisfy Eq. (98).

To prove the first part, first consider, an eigenvector of $\mathbf{U}$:

$$\mathbf{U}\mathbf{u}_i = \lambda_i \mathbf{u}_i \quad . \tag{99}$$

By Theorem 1, λ_i is constant. Hence, differentiating, we obtain

$$\frac{d\mathbf{U}}{dz}\mathbf{u}_i + \mathbf{U}\frac{d\mathbf{u}_i}{dz} = \lambda_i \frac{d\mathbf{u}_i}{dz}$$

or

$$\begin{aligned} \frac{d\mathbf{U}}{dz}\mathbf{u}_i &= j\mathbf{U}\mathbf{V}\mathbf{u}_i - j\mathbf{V}\lambda_i\mathbf{u}_i \\ &= j\mathbf{U}\mathbf{V}\mathbf{u}_i - j\mathbf{V}\mathbf{U}\mathbf{u}_i \\ &= -j[\mathbf{V}, \mathbf{U}]\mathbf{u}_i \quad . \end{aligned} \tag{100}$$

Consider, now, a generalized eigenvector such that

$$\mathbf{U}\mathbf{u}_i = \lambda_i\mathbf{u}_i + \mathbf{u}_{i-1} \quad . \tag{101}$$

Again differentiating

$$\frac{d\mathbf{U}}{dz}\mathbf{u}_i = -\mathbf{U}\frac{d\mathbf{u}_i}{dz} + \lambda_i\frac{d\mathbf{u}_i}{dz} + \frac{d\mathbf{u}_{i-1}}{dz}$$

$$= j\mathbf{UVu}_i - j\mathbf{V}(\lambda_i\mathbf{u}_i + \mathbf{u}_{i-1})$$

$$= j\mathbf{UVu}_i - j\mathbf{VUu}_i$$

$$= -j[\mathbf{V}, \mathbf{U}]\mathbf{u}_i \quad . \tag{102}$$

Hence, Eq. (100) holds for any $\mathbf{u}_i$, eigenvector or generalized eigenvector. Since it holds for a complete set of eigenvectors, the operators must be equal and Eq. (95) must follow.

To prove the second part, we again consider first an eigenvector and differentiate Eq. (99), substituting now from Eq. (95). We find

$$(\mathbf{U} - \lambda_i\mathbf{I})\left(\frac{d\mathbf{u}_i}{dz} + j\mathbf{Vu}_i\right) = 0 \quad . \tag{103}$$

Alternatively, if we consider a generalized eigenvector and differentiate Eq. (101), we find

$$(\mathbf{U} - \lambda_i\mathbf{I})\left(\frac{d\mathbf{u}_i}{dz} + j\mathbf{Vu}_i\right) = \frac{d\mathbf{u}_{i-1}}{dz} + j\mathbf{Vu}_{i-1} \quad . \tag{104}$$

Equation (103) is an eigenvector relation and Eq. (104) is a generalized eigenvector relation. That is, for example, Eq. (103) states that the vector $(d\mathbf{u}_i/dz + j\mathbf{Vu}_i)$ must be in the subspace spanned by the eigenvectors of $\mathbf{U}$ with eigenvalues λ_i.

If there is degeneracy, we can so resolve the degeneracy that $(d\mathbf{u}_i/dz + j\mathbf{Vu}_i)$ is a scalar function times $\mathbf{u}_i$

$$\frac{d\mathbf{u}_i}{dz} + j\mathbf{Vu}_i = f_i(z)\mathbf{u}_i \tag{105}$$

whether $\mathbf{u}_i$ is an eigenvector or a generalized eigenvector.

Let us now define a new set of vectors by the relation

$$\mathbf{u}_i = e^{j\phi_i}\mathbf{v}_i \quad . \tag{106}$$

If we let

$$\phi_i = -j\int_0^z f_i(z)dz \tag{107}$$

then the right side of Eq. (105) vanishes and we are left with

$$\frac{d\mathbf{v}_i}{dz} = -j\mathbf{V}\mathbf{v}_i \quad . \tag{108}$$

Hence, Eq. (95) does imply that it is possible to choose the set of eigenvectors and generalized eigenvectors $\{\mathbf{u}_i\}$ so that Eq. (98) is true.

We note that we have a considerable variation allowed to us, both in the set $\{\mathbf{u}_i\}$ and in the corresponding matrix **V**. As regards the set of vectors, any degeneracies may be resolved arbitrarily, and the manner of resolution may be itself a function of z. Each eigenvector may be multiplied by an arbitrary non-vanishing scalar function of z without affecting its being an eigenvector. But to each such set of vectors there is a particular **V**.

Conversely, we can, in Eq. (95) add to **V** any matrix which commutes with **U** without essentially changing Eq. (95). In particular, we can add any suitable function of **U** [a function of $f(x, z)$ is suitable if for all z, it is analytic with respect to x in a region that includes all the eigenvalues of **U**. We may then add to **V** the matrix $f(\mathbf{U}, z)$.] The corresponding set of vectors is, then, multiplied by appropriate scalar functions.

We must, therefore, study how we can conveniently specialize **V**. The following theorem begins this process.

Theorem 3—If **U** is K-hermitian at all z, then **V** may be chosen to be K-hermitian also at all z.

This follows quickly from the fact that any matrix may be separated into two components:

$$\mathbf{V} = \mathbf{V}_1 + j\mathbf{V}_2 \tag{109}$$

$$\mathbf{V}_1 = \frac{1}{2}(\mathbf{V} + \mathbf{K}^{-1}\mathbf{V}^\dagger\mathbf{K}) \tag{110}$$

$$\mathbf{V}_2 = -\frac{1}{2}j(\mathbf{V} - \mathbf{K}^{-1}\mathbf{V}^\dagger\mathbf{K}) \tag{111}$$

where $\mathbf{V}_1$ and $\mathbf{V}_2$ are both *K*-hermitian.

Further, if $\mathbf{U}$ is everywhere *K*-hermitian, so is $d\mathbf{U}/dz$, as may be seen by differentiating Eq. (87), and remembering that $\mathbf{K}$ is constant. Hence $-j[\mathbf{V}, \mathbf{U}]$ is *K*-hermitian, so that

$$\begin{aligned}
&-j\mathbf{K}(\mathbf{V}_1 + j\mathbf{V}_2)\mathbf{U} + j\mathbf{K}\mathbf{U}(\mathbf{V}_1 + j\mathbf{V}_2) \\
&= j\mathbf{U}^\dagger(\mathbf{V}_1{}^\dagger - j\mathbf{V}_2{}^\dagger)\mathbf{K} - j(\mathbf{V}_1{}^\dagger - j\mathbf{V}_2)\mathbf{U}^\dagger\mathbf{K} \\
&\quad -j\mathbf{K}\mathbf{V}_1\mathbf{U} + \mathbf{K}\mathbf{V}_2\mathbf{U} + j\mathbf{K}\mathbf{U}\mathbf{V}_1 - \mathbf{K}\mathbf{U}\mathbf{V}_2 \\
&= j\mathbf{U}^\dagger\mathbf{V}_1{}^\dagger\mathbf{K} + \mathbf{U}^\dagger\mathbf{V}_2{}^\dagger\mathbf{K} - j\mathbf{V}_1{}^\dagger\mathbf{U}^\dagger\mathbf{K} - \mathbf{V}_2{}^\dagger\mathbf{U}^\dagger\mathbf{K} \\
&= j\mathbf{K}\mathbf{U}\mathbf{V}_1 + \mathbf{K}\mathbf{U}\mathbf{V}_2 - j\mathbf{K}\mathbf{V}_1\mathbf{U} - \mathbf{K}\mathbf{V}_2\mathbf{U}
\end{aligned} \tag{112}$$

since $\mathbf{U}$, $\mathbf{V}_1$, and $\mathbf{V}_2$ are all *K*-hermitian.

Hence

$$2\mathbf{K}\mathbf{V}_2\mathbf{U} = 2\mathbf{K}\mathbf{U}\mathbf{V}_2 \tag{113}$$

or $\mathbf{V}_2$ commutes with $\mathbf{U}$.

But we can add to $\mathbf{V}$ any matrix that commutes with $\mathbf{U}$ without changing Eq. (95) in any essential way. Specifically, we can add $-j\mathbf{V}_2$ and be left only with the *K*-hermitian $\mathbf{V}_1$.

We can, without loss of generality, require $\mathbf{V}$ to be everywhere *K*-hermitian if $\mathbf{U}$ is.

This constraint on $\mathbf{V}$ implies a constraint on the eigenvectors of $\mathbf{U}$, as indicated in the following theorem.

Theorem 4--If $\mathbf{U}$ is *K*-hermitian for all z, and if, in Eq. (95), $\mathbf{V}$ is chosen to be *K*-hermitian for all z, then the *K*-products of the corresponding eigenvectors and generalized eigenvectors of **U** are invariant. In particular, it is possible to so choose them so that they are everywhere K-orthogonal and maximally normalized.

Consider the derivative of such a product:

$$\frac{d}{dz}(\mathbf{u}_i^\dagger \mathbf{K}\mathbf{u}_{\sim j}) = \frac{d\mathbf{u}_i^\dagger}{dz}\mathbf{K}\mathbf{u}_{\sim j} + \mathbf{u}_i^\dagger \mathbf{K}\frac{d\mathbf{u}_{\sim j}}{dz}$$

$$= j\mathbf{U}_i^\dagger \mathbf{V}^\dagger \mathbf{u}_{\sim j} - j\mathbf{u}_i^\dagger \mathbf{K}\mathbf{V}\mathbf{u}_{\sim j}$$

$$= -j\mathbf{u}_i^\dagger(\mathbf{K}\mathbf{V} - \mathbf{V}^\dagger\mathbf{K})\mathbf{u}_{\sim j} \tag{114}$$

where we have used Eq. (98) of Theorem 1. If **V** is K-hermitian, the right hand side vanishes and we must have

$$\mathbf{u}_i^\dagger \mathbf{K}\mathbf{u}_{\sim j} = \lambda_{ij} \tag{115}$$

where λ_{ij} is constant, independent of z.

On the other hand since **U** is K-hermitian, we can, for any particular value of z, choose its eigenvectors and generalized eigenvectors so that

$$\gamma_{ij} = \sigma_i \delta_{ij} \tag{116}$$

as stated in Eq. (94). If the set is chosen so that this is true at one value, and if **V** is K-hermitian, we have just shown that it must be true at all values. Hence the set $\{\mathbf{u}_i\}$ can be chosen to be K-orthogonal and maximally normalized at all z.

Conversely, since **U** is assumed everywhere K-hermitian, we can then choose the set $\{\mathbf{u}_i\}$ to be K-orthogonal and maximally normalized everywhere. If we do so, then the **V** corresponding to these vectors will necessarily be K-hermitian everywhere. This follows by differentiating Eq. (94) and substituting Eq. (98). We find that

$$\mathbf{u}_i^\dagger(\mathbf{K}\mathbf{V} - \mathbf{V}^\dagger\mathbf{K})\mathbf{u}_{\sim j} = 0 \tag{117}$$

for all i and j and at all z. Since the set is complete, we must have **V** K-hermitian everywhere.

This property of K-orthogonality and invariant maximal normalization permits us to write $\mathbf{V}$ explicitly as

$$\mathbf{V} = j \sum_i \sigma_i \frac{d\mathbf{u}_i}{dz} \mathbf{u}_{\sim i}{}^{\dagger}\mathbf{K} \quad . \tag{118}$$

We can do this formally by postmultiplying Eq. (14) by $\sigma_i \mathbf{u}_{\sim i}{}^{\dagger}\mathbf{K}$ and adding over all i. Since, for a K-orthogonal and maximally normalized set

$$\sum_i \sigma_i \mathbf{u}_i \mathbf{u}_{\sim i}{}^{\dagger}\mathbf{K} = \mathbf{I} \tag{119}$$

Equation (118) follows.

Since $\mathbf{V}$ is K-hermitian, we can, alternatively, write

$$\mathbf{V} = \mathbf{K}^{-1}\mathbf{V}^{\dagger}\mathbf{K} = -j \sum \sigma_i \mathbf{u}_{\sim i} \frac{d\mathbf{u}_i{}^{\dagger}}{dz} \mathbf{K} \tag{120}$$

or, perhaps better since it explicitly shows $\mathbf{V}$ to be K-hermitian,

$$\mathbf{V} = \frac{1}{2} j \sum \sigma_i \left(\frac{d\mathbf{u}_i}{dz} \mathbf{u}_{\sim i}{}^{\dagger}\mathbf{K} - \mathbf{u}_i \frac{d\mathbf{u}_{\sim i}{}^{\dagger}}{dz} \mathbf{K} \right) . \tag{121}$$

We should also note that we are still very far from a complete specification of either the set $\{\mathbf{u}_i\}$ or of $\mathbf{V}$. Regarding $\mathbf{V}$, we can still add to it any matrix function of z that is K-hermitian and that commutes with $\mathbf{U}$. In particular we can add any suitable *real* function of $\mathbf{U}$.

Regarding the set $\{\mathbf{u}_i\}$, we can make a substitution of the form of Eq. (106), where the ϕ_i are scalar functions of z, without altering either the orthogonality or normalization properties providing only that

$$\phi_{\sim i} = \phi_i^{*} \quad . \tag{122}$$

We have, still, an enormous amount of freedom available to us. We would like to be able to use this remaining freedom to assure that the resultant $\mathbf{V}$ is expressible by an equation similar to Eq. (95)—*i.e.*, so that the resultant $\mathbf{V}$ has constant eigenvalues. We have not so far found a general way to do this.

We can push the problem somewhat further, however. If we differentiate Eq. (121), we obtain

$$\frac{d\mathbf{V}}{dz} = \frac{1}{2} j \sum \sigma_i \left(\frac{d^2 \mathbf{u}_i}{dz^2} \mathbf{u}_{\sim i} \mathbf{K} - \mathbf{u}_i \frac{d^2 \mathbf{u}_{\sim i}{}^{\dagger}}{dz^2} \mathbf{K} \right) \quad . \tag{123}$$

If, now, it is possible to define a matrix $\mathbf{W}$ so that

$$\frac{d^2 \mathbf{u}_i}{dz^2} = -2j\mathbf{W} \frac{d\mathbf{u}_i}{dz} \tag{124}$$

where $\mathbf{W}$ is K-hermitian, then we can substitute this in Eq. (123) and use Eqs. (118) and (120) to give

$$\frac{d\mathbf{V}}{dz} = -j\,[\mathbf{W}, \mathbf{V}] \quad . \tag{125}$$

The problem, therefore, is to so specify the $\{\mathbf{u}_i\}$ after a transformation of the form of Eq. (106), so that $\mathbf{W}$, defined by Eq. (124) is K-hermitian. This is a problem that will require further study.

This, then completes the general development of the technique of analysis in its present state. Before considering a specific example, we will discuss briefly the technique of obtaining solutions in those cases in which the sequence of Eqs. (85), (86), (90), etc., is short.

D. SOLUTION OF THE SIMPLER CASES

1. $\mathbf{R} = \mathbf{R}_0$, Constant

The simplest of all systems of this sort is when $\mathbf{R}$ itself is constant, $\mathbf{R}_0$. This is the uniform case, and has the formally simple solution

$$\mathbf{x}(z) = e^{-j\mathbf{R}_0 z}\mathbf{x}(0) \quad . \tag{126}$$

There may be a practical difficulty in obtaining the exponential in convenient form, but, at least in principle, it can always be done.

2. $\mathbf{A} = \mathbf{A}_0$, Constant

The next case is where the $\mathbf{A}$ of Eq. (86) is constant. The solution to Eq. (86) is, then

$$\mathbf{R} = e^{-j\mathbf{A}_0 z}\mathbf{R}_0 e^{j\mathbf{A}_0 z} \tag{127}$$

where $\mathbf{R}_0$ is $\mathbf{R}(0)$. Again there is no difficulty, at least in principle, since $\mathbf{A}_0$ commutes with the exponential, and the derivative can be taken without difficulty.

To solve Eq. (85), then, we let

$$\mathbf{x}(z) = e^{-j\mathbf{A}_0 z}\mathbf{y}(z) \tag{128}$$

so that Eq. (85) becomes

$$-je^{-j\mathbf{A}_0 z}\mathbf{A}_0\mathbf{y} + e^{-j\mathbf{A}_0 z}\frac{d\mathbf{y}}{dz} = -je^{-j\mathbf{A}_0 z}\mathbf{R}_0\mathbf{y}$$

or

$$\frac{d\mathbf{y}}{dz} = -j(\mathbf{R}_0 - \mathbf{A}_0)\mathbf{y} \tag{129}$$

and

$$\mathbf{y}(z) = e^{-j(\mathbf{R}_0-\mathbf{A}_0)z}\mathbf{y}(0) \quad . \tag{130}$$

Hence we find that

$$\mathbf{x}(z) = e^{-j\mathbf{A}_0 z}e^{-j(\mathbf{B}_0-\mathbf{A}_0)z}\mathbf{x}(0) \quad . \tag{131}$$

[This form is different from that given in Part II-A of Interim Report 2, Eq. (97). It may be shown to be equivalent, however, and appears to be a more convenient form.]

3. $\mathbf{B} = \mathbf{B}_0$, Constant

The next simplest case is when the $\mathbf{B}$ of Eq. (90) is a constant, $\mathbf{B}_0$.

As above, we have, then,

$$\mathbf{A} = e^{-j\mathbf{B}_0 z}\mathbf{A}_0 e^{j\mathbf{B}_0 z} \quad . \tag{132}$$

If, in Eq. (86), we let

$$\mathbf{R} = e^{-j\mathbf{B}_0 z}\mathbf{S}e^{j\mathbf{B}_0 z} \tag{133}$$

and note that $\mathbf{B}_0$ commutes with its exponential so that

$$\frac{d}{dz}(e^{-j\mathbf{B}_0 z}) = -j\mathbf{B}_0 e^{-j\mathbf{B}_0 z} = -je^{-j\mathbf{B}_0 z}\mathbf{B}_0 \quad .$$

Then Eq. (86) becomes

$$\frac{d\mathbf{S}}{dz} = -j(\mathbf{B}_0 - \mathbf{A}_0)\mathbf{S} + j\mathbf{S}(\mathbf{A}_0 - \mathbf{B}_0) \quad . \tag{134}$$

Hence we find that

$$\mathbf{S} = e^{-j(\mathbf{A}_0-\mathbf{B}_0)z}\mathbf{S}_0 e^{j(\mathbf{A}_0-\mathbf{B}_0)z} \tag{135}$$

and

$$\mathbf{R} = e^{-j\mathbf{B}_0 z}e^{j(\mathbf{A}_0-\mathbf{B}_0)z}\mathbf{R}_0 e^{j(\mathbf{A}_0-\mathbf{B}_0)z}\ e^{j\mathbf{B}_0 z} \tag{136}$$

If, now, in Eq. (85) we let

$$\mathbf{x} = e^{-j\mathbf{B}_0 z}e^{-j(\mathbf{A}_0-\mathbf{B}_0)z}\mathbf{y} \tag{137}$$

we obtain

$$\frac{d\mathbf{y}}{dz} = -j\mathbf{R}_0\mathbf{y} + je^{j(\mathbf{A}_0-\mathbf{B}_0)z}\mathbf{A}_0 e^{-j(\mathbf{A}_0-\mathbf{B}_0)z}\mathbf{y} \quad . \tag{138}$$

This expression can be solved by the lemma that, if **M** and **N** are constant matrices, then

$$e^{\mathbf{M}z}\mathbf{N}e^{-\mathbf{M}z} = \sum_{n=0}^{\infty}[_n\mathbf{M},\ \mathbf{N}]z^n/n! \tag{139}$$

where

$[_k\mathbf{M},\ \mathbf{N}]$ is the k-commutator

that is, where

$$\begin{aligned}
[_0\mathbf{M},\ \mathbf{N}] &= \mathbf{N} \\
[_1\mathbf{M},\ \mathbf{N}] &= [\mathbf{M},\ \mathbf{N}] \\
[_2\mathbf{M},\ \mathbf{N}] &= [\mathbf{M}[\mathbf{M},\ \mathbf{N}]] \\
[_{k+1}\mathbf{M},\ \mathbf{N}] &= [\mathbf{M}[_k\mathbf{M},\ \mathbf{N}]] \quad .
\end{aligned} \tag{140}$$

Hence

$$\frac{d\mathbf{y}}{dz} = -j\mathbf{R}_0\mathbf{y} + j\mathbf{A}_0 y + j\left(\sum_{n=1}^{\infty}[_n j(\mathbf{A}_0 - \mathbf{B}_0),\ \mathbf{A}_0]z^n/n!\right)\mathbf{y}$$

$$= -j(\mathbf{R}_0 - \mathbf{A}_0) + j\sum_{n=1}^{\infty}\mathbf{C}_n\frac{z^n}{n!}\mathbf{y} \tag{141}$$

where

$$\mathbf{C}_n = [_n j(\mathbf{A}_0 - \mathbf{B}_0),\ \mathbf{A}_0] \quad . \tag{142}$$

Then

$$\mathbf{y} = e^{-j\{(\mathbf{R}_0-\mathbf{A}_0)z-\Sigma\mathbf{C}_n z^{n+1}/ n+1)!\}}\mathbf{y}_0 \tag{143}$$

and $\mathbf{x}$ is obtained as

$$\mathbf{x} = e^{-j\mathbf{B}_0 z}e^{-j(\mathbf{A}_0-\mathbf{B}_0)z}e^{-j\{(\mathbf{R}_0-\mathbf{A}_0)z-\Sigma\mathbf{C}_n z^{n+1}/(n+1)!\}}\mathbf{x}_0 \tag{144}$$

The complexity of this solution is formidable. Nevertheless it is an exact solution to a problem that admits a considerable degree of non-uniformity.

We shall not attempt, at this time, the exact solution of the next higher class of problems.

E. NON-UNIFORM TRANSMISSION LINE

We shall, now, consider a specific example: a transmission line whose impedance is a specific function of z.

This problem is one of general importance, primarily because of the paucity of known soluble cases in spite of the importance to microwave engineering of various impedance-transforming or matching transmission-line sections. Our concern here, however, is to determine the significance of the various operations we have performed, and to gain insight into the nature of the physical problems to which this analytic technique applies.

In a lossless transmission line (single mode), described in the E-I basis, the $\mathbf{R}$ matrix is

$$\mathbf{R} = \begin{pmatrix} 0 & \beta Z \\ \beta/Z & 0 \end{pmatrix} \tag{145}$$

where β is the propagation constant and Z is the characteristic equation.

We shall assume β to be constant and allow Z to vary with z. Since the eigenvalues of $\mathbf{R}$ are $\pm\beta$, this puts us in position to apply the techniques developed here.

It may be seen that a suitable metric for this system, independent of z, is

$$\mathbf{K} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{146}$$

which may be found to lead to the conservation of energy.

We could consider the eigenvectors of Eq. (145). However, we might as well consider a somewhat more general situation. It is not difficult to determine that the general matrix $\mathbf{M}(z)$ which is K-hermitian and which has constant eigenvalues is given by

$$\mathbf{M}(z) = \begin{pmatrix} a + jf & g \\ h & a - jf \end{pmatrix} \tag{147}$$

where a is a real constant and f, g, h are real functions of z such that

$$f^2 - gh = b - a^2 = c \quad . \tag{148}$$

Here, b is a real constant. (The constancy of the eigenvalues requires that, in the characteristic equation, the coefficient of each power of λ shall be constant.)

We could, now, proceed by finding the eigenvectors and going through the full process. In the present case involving only 2×2 matrix, it is easier to use brute force. We seek an $\mathbf{N}$ of the form similar to Eq. (147)

$$\mathbf{N} = \begin{pmatrix} A + jF & G \\ H & A - jF \end{pmatrix} \tag{149}$$

such that

$$\frac{d\mathbf{M}}{dz} = -j\,[\mathbf{N},\ \mathbf{M}] \quad . \tag{150}$$

We find that we require

$$f' = Hg - Gh$$

$$g' = 2Fg - 2Gh$$

$$h' = 2Hf - 2Fh \quad . \tag{151}$$

These equations, together with the analogue of Eq. (146)

$$F^2 - GH = C \tag{152}$$

where C is a real constant, determine the possible forms of F, G, and H. They are consistent since, if we substitute them in the derivative of Eq. (146)

$$2ff' - gh' - hg' = 0$$

we do find that the equation vanishes identically.

We could solve Eq. (151) in terms of any one of the functions. However, to do so would imply the non-vanishing of various terms, which we would rather not do.

Let us consider now, our original problem. To put the **R** of of Eq. (145) in the form of **M**, we set

$$g = \beta Z$$

$$h = \beta / Z$$

$$f = 0 \quad . \tag{153}$$

If we substitute these into Eq. (151) we find that we can write:

$$\mathbf{A} = \begin{pmatrix} j\kappa & wZ \\ w/Z & -j\kappa \end{pmatrix} \tag{154}$$

where

$$\kappa = \frac{Z'}{2Z} = \frac{d}{dz} \ln Z^{1/2} \tag{155}$$

and w is an arbitrary real function of z. The condition of Eq. (152) requires that w be of the form

$$w = (\kappa^2 - c)^{\frac{1}{2}} \tag{156}$$

where c is an arbitrary constant, restricted only to being not greater than the minimum of κ^2.

The matrix $\mathbf{A}$ is constant only if κ is constant and we set $c = \kappa^2$. By Eq. (155) this is the case of an exponentially tapered impedance, which is too well known to be interesting.

If, now, we consider the next order of complexity, we find that we must satisfy

$$\begin{aligned} w(ZH) - w\left(\frac{G}{Z}\right) &= \kappa' \\ 2wF - 2\kappa\left(\frac{G}{Z}\right) &= w' + 2w\kappa \\ -2wF + 2\kappa(ZH) &= w' - 2w\kappa \quad . \end{aligned} \tag{157}$$

These equations are consistent. If we subtract 2κ times the first equation from w times the sum of the second and third, we are left with

$$\kappa\kappa' = ww' \tag{158}$$

and this is true by Eq. (156) which holds.

The simplest solution to Eq. (157) is obtained by setting $G = 0$. Then

$$\begin{aligned} H &= \frac{\kappa'}{wZ} \\ F &= \frac{w'}{2w} + \kappa \quad . \end{aligned} \tag{159}$$

For $\mathbf{B}$ to be constant, F must be constant. Substituting from Eq. (156) for w', we must have

$$\frac{\kappa\kappa'}{(\kappa^2 - c)} + \kappa = c'$$

or

$$\frac{\kappa d\kappa}{(c' - \kappa)(c - \kappa^2)} = -dz \qquad (160)$$

or,

$$\frac{-c'}{(c - c^{1/2})} ln\ (c' - \kappa) - \frac{1}{2(c^{1/2} + c')} ln\ (c^{1/2} + \kappa)$$

$$+ \frac{1}{2(c^{1/2} - c')} ln\ (c^{1/2} - \kappa) = a - z \quad . \qquad (161)$$

If, on the other hand, we set $c^{1/2} = c'$, then the solution is

$$\frac{1/2}{(c' - \kappa)} + \frac{1}{4c'} ln\ \frac{c' - \kappa}{c' + \kappa} = a - z \quad . \qquad (162)$$

The required variations of impedance are found by integrating κ to give $ln\ Z^{1/2}$.

Other solutions are possible, which remain to be worked out. These solutions are sufficient, however, to demonstrate that there is a class of impedance variations that are soluble, since they give a constant $\mathbf{B}$ in Eq. (90).

We have shown that there does exist a class of non-uniform systems that are exactly soluble. We have shown some of the properties of these systems, and have opened the way to their more detailed analysis.

There remain some serious questions of a fundamental nature. These were outlined in the introductory section. We have not yet explored the specific case of a non-uniform transmission line in sufficient depth to give an adequate "feel" of the significance of the various situations.

Thus there remains a considerable amount of additional investigation that is needed. However, the approach does seem to be important in that it does provide means for elucidating some of the phenomena that are possible in non-uniform systems.

III PROGRAM FOR THE NEXT INTERVAL

During the next interval we will continue work on the following items:

(1) The multiple pumping situation described in Sec. II-A of the First Interim Report will be considered further. We will attempt to determine, at least in our own minds, if this leads to a practical design of a potentially useful device.

(2) The analysis of commutator-derived systems given in Sec. II will be continued. In particular, we shall look for ways to determine the gross properties of such systems so that the implications of this analysis can be interpreted in terms of physical behavior and device possibilities.

(3) The First Technical Note on the topological analysis of pairwise coupled systems will be issued.

(4) Further study will be made of the group theoretic approach, if time permits, to determine what use can be made of it.

(5) Surveillance of the current literature will be continued.

APPENDIX

THE ELIMINATION OF SELF-COUPLING TERMS IN R

APPENDIX

THE ELIMINATION OF SELF-COUPLING TERMS IN R

We shall here show that it is always possible to remove the self-coupling terms from a representation of the system operator. That is, if the system is described by

$$\frac{d\mathbf{x}}{dz} = -j\mathbf{R}\mathbf{x} \tag{A-1}$$

and $\mathbf{R}$ is expanded in terms of a set of constant K-orthogonal maximally normalized vectors $\{\mathbf{u}_i\}$

$$\mathbf{R} = \sum_{ij} r_{ij}\sigma_i \mathbf{u}_i \mathbf{u}_{\sim j}{}^{\dagger}\mathbf{K} \tag{A-2}$$

then it is always possible to eliminate the terms in $\mathbf{u}_i\mathbf{u}_{\sim i}{}^{\dagger}\mathbf{K}$. We call these the "self-coupling" terms since they produce a change in the $\mathbf{u}_i$ component of x that is in the direction of $\mathbf{u}_i$.

To simplify the symbology somewhat, let us set

$$\mathbf{E}_{ij} = \sigma_i \mathbf{u}_i \mathbf{u}_{\sim j}{}^{\dagger}\mathbf{K} \quad . \tag{A-3}$$

It will be noted that the $\mathbf{E}_{ii}$ are, then, idempotent, and the $\mathbf{E}_{ij}(j \neq i)$ are nilpotent. More generally:

$$\mathbf{E}_{ij}\mathbf{E}_{kh} = \delta_{jk}\mathbf{E}_{ih} \quad . \tag{A-4}$$

Then Eq. (A-2) can be written

$$\mathbf{R} = \sum_{ij} r_{ij}(z)\mathbf{E}_{ij} \quad . \tag{A-5}$$

The elimination from $\mathbf{R}$ of the self-coupling terms then depends on the following lemma:

Lemma 1: If α_k is a scalar and $\mathbf{E}_{kk}$ and idempotent dyad, then

$$e^{-j\alpha_k \mathbf{E}_{kk}} = \mathbf{I} - (1 - e^{-j\alpha_k})\mathbf{E}_{kk} \quad . \tag{A-6}$$

This is proved by the series expansion of the exponential, using the idempotency of $\mathbf{E}_{kk}$:

$$\begin{aligned} e^{-j\alpha_k \mathbf{E}_{kk}} &= \mathbf{I} - j\alpha_k \mathbf{E}_{kk} - \frac{\alpha_k^2}{2!} \mathbf{E}_{kk}^2 + \ldots \sum_n (-j)^n \frac{\alpha_k^n}{n!} \mathbf{E}_{kk}^n \\ &= \mathbf{I} - \mathbf{E}_{kk} \\ &\quad + 1 - \left(j\alpha_k - \frac{\alpha_k^2}{2!} \ldots \sum_n (-j)^n \frac{\alpha_k^n}{n!} \right) \mathbf{E}_{kk} \\ &= \mathbf{I} - \mathbf{E}_{kk} + e^{j\alpha_k} \mathbf{E}_{kk} \\ &= \mathbf{I} - (1 - e^{-j\alpha_k})\mathbf{E}_{kk} \quad . \end{aligned}$$

As an aside which we shall not use here, but for possible future reference, we may note the following lemma:

Lemma 2: If α is a scalar and $\mathbf{E}_{ij}$ a nilpotent matrix, then

$$e^{-j\alpha \mathbf{E}_{ij}} = \mathbf{I} - j\alpha \mathbf{E}_{ij} \quad . \tag{A-7}$$

Again, we prove the lemma by expanding the exponential in a power series, this time obtaining the results directly.

Returning now to our immediate purpose, consider a change of variable given by

$$\mathbf{x} = e^{-j\alpha_k \mathbf{E}_{kk}} \mathbf{y} \tag{A-8}$$

where α_k is a function of z to be chosen as desired. We may note that this change affects only the $\mathbf{u}_k$ component of $\mathbf{x}$. If α_k is real it amounts to making the reference phase angle, against which the phase of this component is measured, a function of position. If the α_k is complex, then there is in addition, a z-dependent renormalization of the $\mathbf{u}_k$ component of x.

We may substitute this into Eq. (A-1) with Eq. (A-5) for $\mathbf{R}$ and obtaining the derivative of the exponential from Eq. (A-6)

$$-j\frac{d\alpha_k}{dz}e^{-j\alpha_k\mathbf{E}_{kk}}\mathbf{y} + e^{-j\alpha_k\mathbf{E}_{kk}}\frac{d\mathbf{y}}{dz}$$

$$= -j\sum_{ij} r_{ij}\mathbf{E}_{ij}\{\mathbf{I} - (1 - e^{-j\alpha_k})\mathbf{E}_{kk}\}\mathbf{y}$$

$$= -j\sum_{i} r_{ik}\, e^{-j\alpha_k}\mathbf{E}_{ik}\mathbf{y}$$

$$-j\sum_{\substack{j\neq k\\ i}} r_{ij}\mathbf{E}_{ij}\mathbf{y} \quad . \tag{A-9}$$

If, now, we set

$$\frac{d\alpha_k}{dz} = r_{kk}$$

$$\alpha_k = \int_0^z r_{kk}(z)dz \tag{A-10}$$

then the first term on the left cancels the term in the first sum on the right with $i = k$. There remains the following equation:

$$e^{ij\alpha_k\mathbf{E}_{kk}}\frac{d\mathbf{y}}{dz} = -j\sum_{i\neq k} r_{ik}e^{-j\alpha_k}\mathbf{E}_{ik}$$

$$-j\sum_{\substack{i\neq k\\ i}} {}_{ij}\mathbf{E}_{ij} \quad . \tag{A-11}$$

Premultiplying by $e^{j\alpha_k \mathbf{E}_{kk}}$, expanded by Eq. (A-6), we obtain

$$\frac{d\mathbf{y}}{dz} = -j\{\mathbf{I} - (1 - e^{j\alpha_k})\mathbf{E}_{kk}\} \sum_{i \neq k} r_{ik} e^{-j\alpha_k \mathbf{E}_{ik}} \mathbf{y}$$

$$-j\{\mathbf{I} - (1 - e^{j\alpha_k})\mathbf{E}_{kk}\} \sum_{\substack{j \neq k \\ i}} e_{ij} \mathbf{E}_{ij} \mathbf{y}$$

$$= -j \sum_{i \neq k} r_{ik} e^{-j\alpha_k} \mathbf{E}_{ik} \mathbf{y}$$

$$-j \sum_{j \neq k} r_{kj} e^{j\alpha_k} \mathbf{E}_{kj} \mathbf{y}$$

$$- j \sum_{\substack{i \neq k \\ j \neq k}} r_{ij} \mathbf{E}_{ij} \mathbf{y} \quad . \tag{A-12}$$

If, now, we set

$$\begin{aligned} r'_{ij} &= 0 \text{ if } i = j = k \\ &= r_{ij} e^{-j\alpha_k} \text{ if } i \neq k,\ j = k \\ &= r_{ij} e^{j\alpha_k} \text{ if } i = k,\ j \neq k \\ &= r_{ij} \text{ if } i \neq k,\ j \neq k \end{aligned} \tag{A-13}$$

then we have

$$\frac{d\mathbf{y}}{dz} = -j\Sigma r_{ij} \mathbf{E}_{ij} \mathbf{y} = -j\mathbf{R}'\mathbf{y} \tag{A-14}$$

by comparison to Eq. (A-5).

We have, in fact, eliminated from $\mathbf{R}$ the term in $\mathbf{E}_{kk}$—*i.e.*, the self-coupling kth component of $\mathbf{R}$. The other terms are multiplied by scalar factors if they couple to or from $\mathbf{u}_k$ ($i = k$, or $j = k$), and are unaffected otherwise. In particular, the other self-coupling terms, *i.e.*, those in $\mathbf{E}_{hh}(h \neq k)$, are unaffected.

We can, therefore, proceed in the same manner, with the elimination of the term in $\mathbf{E}_{hh}$ without reintroducing a term in $\mathbf{E}_{kk}$.

Since the dyads $\mathbf{E}_{ii}$ commute, the law of combination of exponentials applies. We can, therefore, accomplish the simultaneous elimination of the self-coupled terms by the substitution

$$\mathbf{x} = \exp - \sum_k \alpha_k \mathbf{E}_{kk} \mathbf{y}$$

where

$$\alpha_k = \int_0^z r_{kk}(z)\,dz \quad .$$

We have, therefore, a procedure for the complete elimination from $\mathbf{R}$ of all self-coupling terms.

LITERATURE REVIEW

1. K. Blotekjer, "Transverse Electron Beam Noise Described by Filamentary Beam Parameters," *J. Appl. Phys.* **33**, 2409-2414 (August 1962).

   Summary—The effect of transverse emission velocities and positional variation (transverse shot noise) on an equivalent filamentary beam is studied. The frequency dependence of the noise in the various beam modes is obtained. It is shown that there is an effect which is similar to the space-charge smoothing of longitudinal shot noise which acts to reduce the effect of positional variation. This effect, however, does not seem sufficient to account for the low noise observed in, for example, the Adler tube. Hence the author concludes that "successful operation of the cyclotron-wave parametric amplifier is largely due to an effect which is not yet understood."

   Comment—A useful analysis of the generation of the noise that is the input to devices using various mode-coupling. Of interest also is the clear result that existing theories are not wholly adequate.

2. D. C. Forster, "Cooling of the Slow Space-Charge Wave with Application to the TWT," *IRE Trans. PGED*-**9**, 449-453 (November 1962).

   Summary—An analysis of the beam-cooling method originally proposed by Sturrock. In this type of device, a fast wave is first cooled by a Kompfner-null coupler. A slow wave is then coupled to the fast wave parametrically. Because of the parity inversion under parametric coupling, the resultant coupling is passive, or, in our terminology, β-coupling. Hence the noise on the slow wave is transferred to the fast wave. The resultant cooled slow wave can then be used in a conventional TWT structure. The analysis given here is based on linear coupled-mode theory, and involves a detailed study of the effects of the various coupling terms that might be expected. The results of a computer calibration are presented, and indicate a possible noise temperature of about 100°K.

Comment—As far as we know, the feasibility of beam cooling by this method has not yet been demonstrated experimentally. This paper demonstrates to a reasonable degree that the difficulty, if there is one, is due to the neglect of higher-order linear terms. What is perhaps needed, instead, is the investigation of the effect in the much more complicated mode structure found recently by Tore Wessel-Berg by studying thick beams. In other words, this paper is interesting, but may be considering the wrong problem.

3, 4. D. L. Bobroff, H. A. Haus, and J. W. Klüver, "On the Small Signal Power Theorems of Electron Beams," *J. Appl. Phys.* **33**, 2832 (1962), and J. W. Klüver "Potential Form of the Small Signal Power Theorem," *J. Appl. Phys.* **33**, 2943 (1962).

Summary—The small signal power theorem for electron beams has been a subject of controversy for about three years. Certain difficulties in the previous works by the authors[1,2] on this subject have been discussed by E. L. Chu.[3-6] The authors have subsequently answered in rebuttal.[7]

The present papers do not appear to clarify the situation. Little material not previously published is included, although the mathematical form of the power theorem has been favorably modified. Instructive examples are presented.

The following is a bibliography of some of the pertinent articles on this subject:

1. H. A. Haus and D. L. Bobroff, "Small Signal Power Theorem for Electron Beams," *J. Appl. Phys.* **28**, 694 (1957).
2. J. W. Klüver, "Small Signal Power Conservation Theorem for Irrotational Electron Beams," *J. Appl. Phys.* **29**, 618 (1958).
3. E. L. Chu, "Two Alternative Definitions of Small Signal RF Power of Electron Beams," *J. Appl. Phys.* **30**, 1617 (1959).
4. E. L. Chu, "Comments on Klüver's Paper Entitled Small Power Conservation Theorem for Irrotational Electron Beams," *J. Appl. Phys.* **30**, 1618 (1959).
5. E. L. Chu, "The Lagrangian and the Energy-Momentum Tensors in the Perturbation Theory of Classical Electrodynamics," *Annals of Phys.* **9**, 76 (1960).

6. E. L. Chu, "On the Concept of Fictitious Surface Charges of an Electron Beam," *J. Appl. Phys.* **31**, 381 (1960).

7. D. L. Bobroff, H. A. Haus, and J. W. Klüver, "On E. L. Chu's Definition of Small-Signal RF Power of Electron Beams," *J. Appl. Phys.* **32**, 749 (1961).

STANFORD RESEARCH INSTITUTE

MENLO PARK CALIFORNIA

Regional Offices and Laboratories

Southern California Laboratories
820 Mission Street
South Pasadena, California

Washington Office
808 17th Street, N.W.
Washington 5, D.C.

New York Office
270 Park Avenue, Room 1770
New York 17, New York

Detroit Office
The Stevens Building
1025 East Maple Road
Birmingham, Michigan

European Office
Pelikanstrasse 37
Zurich 1, Switzerland

Japan Office
911 Iino Building
22, 2-chome, Uchisaiwai-cho, Chiyoda-ku
Tokyo, Japan

Representatives

Honolulu, Hawaii
Finance Factors Building
195 South King Street
Honolulu, Hawaii

London, England
19 Upper Brook Street
London, W. 1, England

Milan, Italy
Via Macedonio Melloni 40
Milano, Italy

London, Ontario, Canada
P.O. Box 782
London, Ontario, Canada